

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 892 047 A2

(12)

EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
20.01.1999 Patentblatt 1999/03

(51) Int. Cl.⁶: C12N 15/12, C07K 14/47,
A61K 48/00, G01N 33/50

(21) Anmeldenummer: 98112470.4

(22) Anmeldetag: 06.07.1998

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 09.07.1997 DE 19729211
11.02.1998 DE 19805371

(71) Anmelder:
Hoechst Marion Roussel Deutschland GmbH
65929 Frankfurt am Main (DE)

(72) Erfinder:
• Fleckenstein, Bernhard, Prof. Dr.
91369 Wiesenthau (DE)
• Ensser, Armin, Dr.
90419 Nürnberg (DE)

(54) Humanes und murines Semaphorin L

(57) Humanes Semaphorin L(H-SemaL) und korrespondierende Semaphorine in anderen Spezies.

Gegenstand der Erfindung sind neuen Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Verwendung derselben.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunoglobulin-ähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

EP 0 892 047 A2

Beschreibung

Humanes Semaphorin L(H-SemaL) und korrespondierende Semaphorine in anderen Spezies.

Gegenstand der Erfindung sind neue Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Herstellung und Verwendung derselben.

Semaphorine wurden erstmals von Koldokin (Koldokin et al. (1993) Cell 75:1389-1399) als Mitglieder einer konservierten Genfamilie beschrieben.

Inzwischen wurden die Gene bzw. Teile der Gene weiterer Semaphorine kloniert und teilweise charakterisiert. Bis-her waren insgesamt 5 humane (H-Sema III, H-Sema V, H-Sema IV, H-SemaB und H-SemaE) (Koldokin et al. (1993); Roche et al. (1996) Onkogene 12:1289-1297; Sekido et al. (1996) Proc. Natl. Acad. Sci. USA 93:4120-4125; Xiang et al. (1996) Genomics 32:39-48; Hall et al. (1996) Proc. Natl. Acad. Sci. USA 93:11780-11785; Yamada et al. (1997) (GenBank Zuordnungs-Nr. AB000220), 8 murine (Gene der Maus; M-Sema A bis M-SemaH) (Püschel et al. (1995) Neuron 14:941-948; Messerschmidt et al. (1995) Neuron 14:949-959; Inigaki et al. (1995) FEBS Letters 370:269-272; Adams et al. (1996) Mech. Dev. 57:33-45; Christensen et al. (1996) (Genbank Zuordnungs-Nr. Z80941, Z93948)), 5 galline (Huhn) (Collapsin-1 bis -5) (Luo et al. (1993); Luo et al. (1995) Neuron 14:1131-1140), und Gene von Ratte (R-Sema III) (Giger et al. (1996) J. Comp. Neurol. 375:378-392), Zebrafisch, Insekten (Fruchtfliege (Drosophila melanogaster: D-Sema I und D-Sema II), Käfer (Tribolium confusum: T-Sema I), Grasshüpfer (Schistocerca americana: G-Sema I)) (Koldokin et al. (1993)), und Nematoden (C.elegans: Ce-Sema) (Roy et al. (1994) (GenBank Zuordnungs-Nr. U15667)) bekannt. Weiterhin besitzen zwei Poxviren (Vaccinia (ORF-A39) und Variola (ORF-A39-homolog)) (Koldokin et al. (1993)) sowie der Alcelaphine Herpesvirus Typ 1 (AHV-1) (AHV-Sema) (Ensser und Fleckenstein (1995) Gen. Virol. 76:1063-1067) zu Semaphorinen homologe Gene.

Einen Überblick über die bisher identifizierten Semaphorine in verschiedenen Spezies gibt Tabelle 1. In Tabelle 1 sind die Namen der Semaphorine (Spalte 1), die verwendeten Synonyme (Spalte 2), die Spezies aus der das jeweilige Semaphorin isoliert wurde (Spalte 3) sowie, soweit bekannt, Daten zur Domänenstruktur des kodierten Proteins und zur chromosomal Lokalisation (Spalte 4 in Tabelle 1), die Zuordnungsnummer unter der die Sequenz des Gens in Gendatenbanken, z.B. in einer EST (expressed sequence tags) Datenbank, EMBL (European Molecular Biology Laboratory, Heidelberg) oder NCBI (National Center for Biotechnology Information, Maryland, USA) gespeichert ist und die entsprechende Referenz unter der diese Daten publiziert wurden (Spalte 5 in Tabelle 1), angegeben.

Alle Genprodukte (kodierte Semaphorine) der bisher bekannten Semaphorin Gene weisen ein N-terminales Signalpeptid auf, an dessen C-terminalem Ende sich eine charakteristische Sema-Domäne mit einer Länge von etwa 450 bis 500 Aminosäuren befindet. Innerhalb der Sema-Domäne finden sich stark konservierte Aminosäuremotive und eine Anzahl hochkonservierter Cysteinreste. Die Genprodukte (Semaphorine) unterscheiden sich in den auf die Sema-Domäne folgenden C-terminalen Sequenzen, die aus einer oder mehreren Domänen aufgebaut sind. Sie weisen beispielsweise in diesen C-terminalen Aminosäuresequenzen Transmembrandomänen (TM), Immunoglobulin-ähnliche Domänen (Ig) (konstanter Teil des Immunoglobulins), zytoplasmatische Sequenzen (CP), Prozessierungssignale (P) (beispielsweise mit der Konsensussequenz (RXR), wobei R für die Aminosäure Arginin und X für eine beliebige Aminosäure steht) und/oder hydrophile C-Termini (HPC) auf. Auf der Basis der unterschiedlichen Domänenstruktur im C-Terminus lassen sich die bisher bekannten Semaphorine in 5 verschiedene Untergruppen einteilen (I bis V):

40	I	Sezerniert, ohne weitere Domäne (z.B. ORF-A49)
	II	Ig Sezerniert (ohne Transmembrandomäne) (z.B. AHV-Sema)
	III	Ig, TM, CP Membranverankert mit zytoplasmatischer Sequenz (z.B. CD100)
	IV	Ig, (P), HPC Sezerniert mit hydrophilem C-Terminus (z.B. H-Sema III, M-SemaD, Collapsin-1)
45	V	Ig, TM, CP Membranverankert mit C-terminalem 7 Thrombospondin-Motiv (z.B. M-SemaF und G)

Ein Rezeptor oder extrazellulärer Ligand für Semaphorine wurde bisher nicht beschrieben. Im Zusammenhang mit Semaphorin-vermittelten Effekten wurden intrazelluläre, heterotrimeren GTP-bindende Proteinkomplexe beschrieben. Als ein Bestandteil dieser Proteinkomplexe wurden bei Hühnern sogenannte CRMP-Proteine (Collapsin response mediator protein) identifiziert, welche vermutlich Bestandteil der Semaphorin-induzierten intrazellulären Signalkaskade sind (Goshima et al. (1995) Nature 376: 509-514). Das CRMP62 beispielsweise besitzt: Homologie zu unc-33, einem für das gerichtete Axonwachstum essentiellen Nematoden-Protein. Ein humanes Protein mit 98% Aminosäure-Identität zu CRMP62 ist ebenfalls bekannt (Hamajima et al. (1996) Gene 180: 157-163. In Ratten wurden ebenfalls mehrere CRMP-verwandte Gene beschrieben (Wang et al. (1996) Neurosci. 16: 6197-6207).

55 Die sezernierten oder transmembranen Semaphorine vermitteln repulsive Signale für wachsende Nervenknospen. Sie spielen eine Rolle bei der Entwicklung des zentralen Nervensystems (ZNS) und werden vor allem in Muskel- und Nervengewebe exprimiert (Koldokin et al. (1993); Luo et al. (1993) Cell 75:217-227).

Außer im ZNS konnte eine deutliche Expression von M-SemaG auch auf Zellen des lymphatischen und hämato-

poetischen Systems beobachtet werden, im Gegensatz zum nahe verwandten M-SemaF (Furuyima et al. (1996) J. Biol. Chem. 271: 33376-33381).

Kürzlich wurden zwei weitere humane Semaphorine identifiziert, H-Sema IV und H-Sema V und zwar in einer Region auf Chromosom 3p21.3, deren Deletion mit verschiedenen Formen von Bronchialkarzinomen assoziiert ist H-Sema IV (Roche et al. (1996), Xiang et al. (1996), Sekido et al. (1996)) ist auf Aminosäureebene etwa zu 50% identisch mit M-SemaE, während H-Sema V (Sekido et al. (1996)) das direkte Homolog zu M-Sema-A ist (86% Aminosäureidentität). Da diese Gene (H-Sema IV und V) im Rahmen von DNA-Sequenzierungs-Projekten der deletierten 3p21.3 Loci gefunden wurden, ist die komplexe Intron-Exon-Struktur dieser beiden Gene bekannt. Die beiden Gene werden in verschiedenen neuronalen und nicht neuronalen Geweben exprimiert.

Ebenfalls erst vor kurzem wurde das zelluläre Oberflächenmolekül CD100 (human), exprimiert und induziert auf aktivierten T-Zellen, als Semaphorin identifiziert (ebenfalls in Tabelle 1 aufgeführt). Es unterstützt die Interaktion mit B-Zellen über den Rezeptor CD40 und den entsprechenden Liganden CD40L. CD100 ist ein membranverankertes Glykoprotein-Dimer von 150kd (Kilodalton). Es wurde eine Assoziation des intrazytoplasmatischen C-Terminus von CD100 mit einer noch unbekannten Kinase beschrieben (Hall et al. (1996)). Damit ist CD100 das erste und bisher einzige Semaphorin, dessen Expression in Zellen des Immunsystems nachgewiesen werden konnte.

Unter der Fragestellung "Transformierende Gene von Rhadinoviren" wurde das komplette Genom des Alcelaphinen Herpesvirus Typ 1 (AHV-1) kloniert und sequenziert (Ensser et al. (1995)). AHV-1 ist der Erreger des bösartigen Katarrhafiebers, einer mit einem lymphoproliferativen Syndrom einhergehenden, meist fatalen Erkrankung verschiedener Wiederkäuer. Bei der Analyse wurde an einem Ende des viralen Genoms ein offener Leserahmen mit entfernter, aber signifikanter Homologie zu einem Gen von Vacciniaivirus (ORF-A39 entspricht VAC-A39 in Ensser et al. (1995)) J. Gen. Virol. 76:1063-1067), welches der Genfamilie der Semaphorine zugerechnet wurde, gefunden. Während das AHV-1 Semaphorin (AHV-Sema) eine gut konservierte Semaphorin-Struktur besitzt, sind die Poxvirus-Gene (ORF-A39 und ORF-A39-homolog, siehe Tabelle 1) C-terminal verkürzt, d.h. bei ihnen ist die konservierte Sema-Domäne nur unvollständig vorhanden.

Ein Datenbankvergleich des gefundenen AHV-Sema mit dbEST (EST (expressed sequence tags)-Datenbank (db)) lieferte jeweils 2 EST-Sequenzen von 2 unabhängigen cDNA-Klonen aus humarer Plazenta (Zuordnungsnummern H02902, H03806 (Klon 151129), Zuordnungsnummern R33439 und R33537 (Klon 135941)). Diese wiesen deutlich höhere Homologie zum AHV-1 Semaphorin auf, als zu den bis dahin beschriebenen neuronalen Semaphorinen.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunglobulinähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

Die Aminosäure-Sequenz des Signalpeptids kann weniger als 70, vorzugsweise weniger als 60 Aminosäuren und mehr als 20, vorzugsweise mehr als 30 Aminosäuren aufweisen, besonders bevorzugt ist eine Länge von etwa 40 bis 50 Aminosäuren. In einer speziellen Ausführungsform der Erfindung hat das Signalpeptid eine Länge von 44 Aminosäuren, d.h. zwischen den Aminosäuren 44 und 45 befindet sich eine Spaltstelle für eine Signalpeptidase.

Die Sema-Domäne kann eine Länge von 300 bis 700 oder mehr, vorzugsweise von etwa 400 bis 600 Aminosäuren aufweisen. Bevorzugt sind Sema-Domänen mit einer Länge von 450 bis 550 Aminosäuren, vorzugsweise von etwa 500 Aminosäuren. In einer besonderen Ausführungsform der Erfindung schließt sich die Sema-Domäne an das Signalpeptid an, wobei sich die Sema-Domäne vorzugsweise bis zur Aminosäure 545 erstreckt.

Die Immunglobulin-ähnliche Domäne kann eine Länge von etwa 30 bis 110 oder mehr Aminosäuren aufweisen, bevorzugt sind Längen zwischen 50 und 90, besonders bevorzugt etwa 70 Aminosäuren.

Die Transmembrandomäne kann eine Länge von etwa 10 bis 35, vorzugsweise von etwa 15 bis 30, besonders bevorzugt von etwa 20 bis 25 Aminosäuren aufweisen.

Gegenstand der Erfindung sind Semaphorine vom Typ L aus verschiedenen Spezies, insbesondere aus Wirbeltieren, beispielsweise aus Vögeln und/oder Fischen, vorzugsweise aus Säugetieren, beispielsweise aus Primaten, Ratte, Kaninchen, Hund, Katze, Schaf, Ziege, Kuh, Pferd, Schwein, besonders bevorzugt aus Mensch und Maus. Gegenstand der Erfindung sind auch entsprechende Semaphorine aus Mikroorganismen, insbesondere aus pathogenen Mikroorganismen, beispielsweise aus Bakterien, Hefen und/oder Viren, z.B. aus Retroviren, insbesondere aus humanpathogenen Mikroorganismen.

Eine Ausführungsform der Erfindung ist ein entsprechendes humanes Semaphorin (H-SemaL), das ein Signalpeptid, eine Sema-Domäne, eine Immunglobulin-ähnliche Domäne und eine Transmembrandomäne aufweist. Eine spezielle Ausführungsform ist das Semaphorin, das durch die Aminosäuresequenz gemäß Tabelle 4 gegeben ist.

Eine weitere Ausführungsform der Erfindung sind korrespondierende Semaphorine in anderen Spezies, die im Bereich der Sema-Domäne eine Aminosäureidentität größer als 40 %, vorzugsweise größer 50 %, besonders bevorzugt größer 60 % im Bezug auf die Sema-Domäne von H-SemaL (Aminosäuren 45 bis 545 der Sequenz in Tabelle 4)

aufweisen. Aus näher verwandten Spezies (z.B. Primaten, Maus) können die korrespondierenden Semaphorine durchaus Aminosäureidentitäten größer als 70%, vorzugsweise größer als 80 %, besonders bevorzugt größer als 90 % aufweisen. Prozentuale Homologien können beispielsweise mit dem Programm GAP (GCG Programm-Paket, Genetic Computer Group (1991)) bestimmt bzw. berechnet werden.

5 Eine derartige Ausführungsform der Erfindung ist ein korrespondierendes Semaphorin der Maus (murines Semaphorin (M-SemaL)). Beispielsweise enthält dieses die partielle Aminosäuresequenz gemäß Tabelle 5 (murines Semaphorin (M-SemaL)).

Die Erfindung betrifft auch korrespondierende Semaphorine, die eine Aminosäureidentität (über die Gesamtlänge der Aminosäuresequenz des Protein betrachtet) von nur etwa 15 bis 20% bei wenig verwandten Spezies (phylogenetisch weit voneinander entfernt), vorzugsweise 25 bis 30%, besonders bevorzugt 35 bis 40 % oder eine höhere Identität im Bezug auf die gesamte Aminosäuresequenz von H-SemaL gemäß Tabelle 4 aufweisen.

10 Die Gene, die für Semaphorine von Typ L kodieren, weisen eine komplexe Exon-Intron-Struktur auf. Diese Gene können beispielsweise zwischen 10 und 20 Exons, vorzugsweise etwa 11 bis 18, besonders bevorzugt 12 bis 16 Exons und eine entsprechende Anzahl von Introns aufweisen. Sie können aber auch die gleiche Anzahl Exons und Introns 15 aufweisen wie das Gen von H-SemaL (13 oder 15 Exons, vorzugsweise 14 Exons). Eine besondere Ausführungsform der Erfindung betrifft das Gen von H-SemaL. Dieses Gen hat vorzugsweise eine Länge von 8888 bis 10000 oder mehr Nukleotiden. Das humane Semaphorin-Gen enthält vorzugsweise die Nukleotid-Sequenz, die in Tabelle 14 gegeben ist oder die Nukleotidsequenz, die in der Datenbank GenBank® unter der Zugangsnummer AF030697 hinterlegt wurde. Diese Nukleotidsequenzen enthalten mindestens 13 Introns. Darüber hinaus weist das humane Semaphorin-Gen am 20 5'-Ende einen zusätzlichen Sequenzbereich auf. Dieser Bereich enthält gegebenenfalls weitere kodierende und nichtkodierende Sequenzen, z.B. ein oder zwei weitere Introns bzw. Exons.

Die Versuche zur chromosomal Lokalisation des humanen Semaphorins vom Typ L ergaben, daß das entsprechende Gen an Position 15q22.3-23 lokalisiert ist. Entsprechend wurde das Gen für M-SemaL an Position 9A3.3-B lokalisiert.

25 Als Folge der komplexen Intron-Exon Struktur kann das Primärtranskript der Semaphorin mRNA unterschiedlich gespliced werden, wodurch unterschiedliche Splicevarianten der Semaphorine entstehen. Die aus diesen Splicevarianten translatierten Proteine sind Derivate der erfindungsgemäßen Semaphorine. Sie entsprechen in ihrer Aminosäuresequenz und auch weitgehend in ihrer Domänenstruktur den beschriebenen, erfindungsgemäßen Semaphorinen vom Typ L, sind jedoch gegebenenfalls gegenüber diesen verkürzt. Beispielsweise können Splicevarianten, denen die 30 Transmembrandomäne ganz oder teilweise fehlt, gebildet werden. Ein Semaphorin-Derivat, welches keine oder keine vollständige Transmembrandomäne, aber ein Signalpeptid enthält, kann sezerniert werden und auf diese Weise außerhalb der Zelle lokal oder auch über größere Entfernung wirken, beispielsweise auf andere Zellen. Eine andere Splicevariante kann beispielsweise keine Sequenz mehr enthalten, die für ein Signalpeptid kodiert und gegebenenfalls auch 35 keine Sequenz, die für eine für hydrophobe Aminosäuresequenz kodiert, die eine potentielle Transmembrandomäne darstellt. Eine Folge wäre, daß dieses Semaphorin-Derivat weder in die Membran eingebaut, noch sezerniert wird (es sei denn über sekretorische Vesikel). Ein solches Semaphorin-Derivat kann an intrazellulären Prozessen, beispielsweise an Signaltransduktionsprozessen beteiligt sein. Auf diese Weise können mit dem gleichen Grundmolekül (Semaphorine vom Typ L) und den davon abgeleiteten Derivaten (beispielsweise Splicevarianten) vielfältige intra- und extrazelluläre Prozesse reguliert und/oder aufeinander abgestimmt werden.

40 Eine besondere Ausführungsform der Erfindung betrifft Semaphorin-Derivate, die sich von den erfindungsgemäßen Semaphorinen vom Typ L ableiten, die aber keine oder keine vollständige Transmembrandomäne enthalten.

Eine weitere Ausführungsform der Erfindung betrifft Semaphorin-Derivate, die sich von den erfindungsgemäßen Semaphorinen vom Typ L ableiten, die aber kein Signalpeptid enthalten.

Das Signalpeptid kann auch posttranslational abgespalten werden. Dadurch wird ein membranständiges (mit TM-

45 Domäne) oder ein sezerniertes (Splicevariante ohne TM-Domäne) Semaphorin-Derivat mit verkürzter Domänenstruktur gebildet. Ein auf diese Weise posttranslational prozessiertes Semaphorin-Derivat enthält nur noch Sema-Domäne, Ig-Domäne und gegebenenfalls Transmembrandomäne. Eine Signalpeptidschnittstelle kann beispielsweise direkt am Ende des Signalpeptids liegen, sie kann z.B. 40 bis 50 Aminosäuren oder weiter vom Aminoterminus entfernt lokalisiert sein.

50 Ein "verkürztes" (d.h. weniger Domänen enthaltendes) Semaphorin L-Derivat ist von anderen Semaphorinen, die sich nicht von den Semaphorinen vom Typ L ableiten, dadurch zu unterscheiden, daß es eine sehr große (> 90 %) Aminosäureidentität oder eine identische Aminosäuresequenz mit den Semaphorinen vom Typ L in den vorhandenen Domänen aufweist.

55 Die erfindungsgemäßen Semaphorine können auch in anderer Weise posttranslational modifiziert sein. Beispielsweise können sie ein-, zwei-, drei-, vier-, fünf-, sechs-, sieben-, acht-, neun-, zehn- oder mehrfach glykosyliert (N- und/oder O-glykosyliert) vorliegen. Die Aminosäuresequenzen der Semaphorine können dann ebenso viele oder mehr Konsensussequenzen für potentielle Glykosylierungsstellen aufweisen, vorzugsweise fünf derartige Stellen. Eine Ausführungsform der Erfindung betrifft Semaphorine, bei denen die Glykosylierungsstellen an Positionen lokalisiert sind,

die den Positionen 105, 157, 258, 330 und 602 der H-SemaL Aminosäuresequenz entsprechen (Tabelle 4).

5 Darüber hinaus können die Semaphorine in Form ihrer phosphorylierten Derivate vorliegen. Semaphorine können die Substrate unterschiedlicher Kinasen sein, beispielsweise können die Aminosäuresequenzen Konsensussequenzen für Protein Kinase C, Tyrosin Kinase und/oder Kreatin Kinase aufweisen. Weiterhin können die Aminosäuresequenzen der Semaphorine Konsensussequenzen für potentielle Myristylierungsstellen aufweisen. An diesen Stellen können entsprechende Semaphorin-Derivate mit Myristinsäure verestert sein.

10 Die erfindungsgemäßen Semaphorine vom Typ L und deren Derivate können in Form von Monomeren, Dimeren und/oder Multimeren vorliegen, beispielsweise können zwei oder mehr Semaphorine bzw. deren Derivate über intermolekulare Disulfidbrücken miteinander verbunden sein. Darüber hinaus können sich intramolekulare Disulfidbrücken ausbilden.

15 Derivate der erfindungsgemäßen Semaphorine sind weiterhin Fusionsproteine. Ein solches Fusionsprotein enthält einerseits ein Semaphorin vom Typ L oder Teile desselben und darüber hinaus ein weiteres Peptid oder Protein bzw. ein Teil desselben. Peptide oder Proteine bzw. Teile derselben können z.B. Epitope-Tags (z.B. His-Tag (6xHistidin), Myc-Tag, flu-Tag), die z.B. zur Aufreinigung der Fusionsproteine verwendet werden können, oder solche, die zur Markierung der Fusionsproteine verwendet werden können, z.B. GFP (green fluorescent protein), sein. Beispiele für Derivate der Semaphorine vom Typ L sind z. B. durch die in den Beispielen beschriebenen Konstrukte gegeben. Die Sequenzen dieser Konstrukte können aus den Tabellen 7 bis 15 gegebenenfalls unter Berücksichtigung der Annotations zu den Plasmiden entnommen werden.

20 Weiterer Gegenstand der Erfindung sind Nukleinsäure-Sequenzen, vorzugsweise DNA- und RNA-Sequenzen, die für die erfindungsgemäßen Semaphorine vom Typ L und/oder deren Derivate kodieren, beispielsweise die entsprechenden Gene, die unterschiedlichen Splicevarianten der mRNA, die dazu korrespondierenden cDNAs sowie deren Derivate, z.B. Salze der DNA bzw. RNA Derivate im Sinne der Erfindungen sind Sequenzen oder Teile davon, die z.B. mit molekularbiologischen Methoden verändert und an die jeweiligen Anforderungen angepaßt werden, beispielsweise verkürzte Gene oder Teile der Gene (z. B. Promotorsequenzen, Terminatorsequenzen), cDNAs oder Chimäre derselben, Konstrukte für Expressionen und Klonierungen und deren Salze.

25 Eine Ausführungsform betrifft die genomischen Sequenzen (Gene) der Semaphorine vom Typ L. Die Erfindung betrifft die Intron- und Exon-Sequenzen und genregulatorische Sequenzen, beispielsweise Promotor-, Enhancer- und Silencer-Sequenzen.

30 Ein Gegenstand dieser Ausführungsform ist das Gen von H-SemaL bzw. dessen Derivate. Ein Gegenstand der Erfindung ist ein Gen, das die Nukleotidsequenz, die in Tabelle 14 gegeben ist, enthält. Ein weiterer Gegenstand der Erfindung ist das Gen, das die Nukleotidsequenz, die in der Datenbank GenBank® unter der Zugangsnummer AF030697 hinterlegt ist, enthält.

35 Ein weiterer Gegenstand dieser Ausführungsform ist das Gen von M-SemaL bzw. dessen Derivate.

40 Ein weiterer Gegenstand der Erfindung ist die cDNA von H-SemaL oder deren Derivate (z. B. Teile der cDNA). Eine besondere Ausführungsform ist die cDNA von H-SemaL gemäß der Nukleotidsequenz in Tabelle 2. Ein weiterer Gegenstand der Erfindung ist die cDNA von H-SemaL, die in der Datenbank GenBank® unter der Zugangsnummer AF030698 hinterlegt ist. Gegenstand der Erfindung sind auch die zu diesen cDNAs korrespondierenden mRNAs bzw. Teile derselben.

45 Ein weiterer Gegenstand der Erfindung ist die cDNA von M-SemaL oder deren Derivate (z. B. Teile der cDNA). Eine besondere Ausführungsform ist die partielle cDNA-Sequenz von M-SemaL gemäß Tabelle 3 sowie cDNA-Sequenzen, die diese partielle cDNA-Sequenz enthalten. Eine weitere Ausführungsform der Erfindung betrifft die cDNA von M-SemaL, die in der Datenbank GenBank unter der Zugangsnummer AF030699 hinterlegt ist. Gegenstand der Erfindung sind auch die zu diesen cDNAs korrespondierenden mRNAs bzw. Teile derselben.

50 Die Erfindung beinhaltet auch Allele und/oder individuelle Ausprägungsformen der Gene/mRNAs/cDNAs, die sich nur geringfügig von den hier beschriebenen Semaphorin-Sequenzen unterscheiden und für ein identisches oder nur geringfügig verändertes Protein (Abweichung in der Aminosäure-Sequenz kleiner oder gleich 10%) kodieren (weiteres Beispiel für Derivate). Weitere Beispiele für die Derivate sind durch die in den Beispielen angegebenen Konstrukte gegeben. Die Sequenzen dieser Konstrukte sind in den Tabellen 7 bis 14 dargestellt und können unter Berücksichtigung der Annotation für Plasmide interpretiert werden.

55 Ein weiterer Gegenstand der Erfindung sind Plasmide, die DNA, die für die Semaphorine des Typs L bzw. deren Derivate kodiert, enthalten. Solche Plasmide können beispielsweise Plasmide mit hohen Replikationsraten sein, die für eine Amplifikation der DNA, z.B. in E. coli geeignet sind.

55 Eine spezielle Ausführungsform sind Expressionsplasmide, mit denen die Semaphorine bzw. Teile davon oder deren Derivate in prokaryotem und/oder eukaryotem Expressions-Systemen exprimiert werden können. Es sind sowohl Expressionsplasmide die konstitutive als auch solche, die induzierbare Promotoren enthalten, geeignet.

55 Gegenstand der Erfindung sind auch Verfahren zur Herstellung von Nukleinsäuren, die für Semaphorine vom Typ L oder Derivate derselben kodieren.

55 Beispielsweise können diese Nukleinsäuren, z. B. DNA oder RNA auf chemischem Weg synthetisiert werden. Insbeson-

dere können diese Nukleinsäuren, z. B. die entsprechenden Gene oder cDNAs bzw. Teile derselben mit der PCR unter Verwendung von spezifischen Primern und geeignetem Ausgangsmaterial als Template (z. B. cDNA aus einem geeigneten Gewebe oder genomische DNA) amplifiziert werden.

Ein konkretes Verfahren zur Herstellung von Semaphorin L cDNA bzw. des H SemaL Gens ist in den Beispielen beschrieben.

5 Die Erfindung betrifft auch Verfahren zur Herstellung der Semaphorine vom Typ L. Beispielsweise kann ein Semaphorin L oder ein Derivat desselben dadurch hergestellt werden, daß eine entsprechende Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L oder ein Derivat desselben kodiert, in einen Expressionsvektor kloniert und mit diesem rekombinanteren Vektor eine geeignete Zelle transformiert wird. Es können beispielsweise prokaryote oder eukaryote 10 Zellen verwendet werden. Die Semaphorine vom Typ L oder deren Derivate können gegebenenfalls auch auf chemischem Weg hergestellt werden.

15 Darüber hinaus können die Semaphorine vom Typ L bzw. deren Derivate als Fusionsproteine exprimiert werden, beispielsweise mit Proteinen oder Peptiden, die einen Nachweis des exprimierten Fusionsproteins erlauben, z.B. als Fusionsprotein mit GFP (green fluorescent protein). Die Semaphorine können auch als Fusionsproteine mit einem, zwei, drei oder mehreren Epitop-Tags, beispielsweise mit Myc- und/oder His-(6xHistidin) und/oder flu-Tags exprimiert werden. Entsprechend können Plasmide verwendet oder hergestellt werden, die DNA-Sequenzen enthalten, die für diese Fusionsproteine kodieren. Beispielsweise können Semaphorin kodierende Sequenzen in Plasmide kloniert werden, die DNA-Sequenzen enthalten, die für GFP und/oder Epitop-Tags, z.B. Myc-Tag, His-Tag, flu-Tag kodieren. Konkrete Beispiele dafür sind durch die Beispiele und die in den Tabellen aufgeführten Sequenzen ggfs. unter 20 Zuhilfenahme der Annotation zu den Plasmiden gegeben.

25 Ein weiterer Gegenstand der Erfindung sind Antikörper, die spezifisch die Semaphorine vom Typ L, deren Derivate oder Teile davon binden bzw. erkennen. Dies können beispielsweise polyklonale oder monoklonale Antikörper, die z.B. in Maus, Kaninchen, Ziege, Schaf, Huhn usw. hergestellt werden können, sein.

30 Eine besondere Ausführungsform dieses Gegenstandes der Erfindung sind Antikörper, die gegen die Epitope die den Aminosäuresequenzen von Position 179 bis 378 bzw. 480 bis 666 der H-SemaL Sequenz gemäß Tabelle 4 entsprechen, gerichtet sind. Die Erfindung betrifft auch ein Verfahren zur Herstellung von spezifischen Anti-Semaphorin L Antikörpern, wobei für die Herstellung Antigene verwendet werden, die genannten Epitope enthalten.

35 Die Erfindung betrifft auch Verfahren zur Herstellung der Antikörper, vorzugsweise wird dazu ein Fusionsprotein, bestehend aus einem charakteristischen Semaphorin Epitop und einem Epitop-Tag, welches für die anschließende Aufreinigung des rekombinannten Fusionsproteins verwendet werden kann. Das aufgereinigte Fusionsprotein kann anschließend für die Immunisierung verwendet werden. Zur Herstellung des rekombinannten Fusionsproteins wird ein entsprechender rekombinanter Expressionsvektor hergestellt und mit diesem eine geeignete Zelle transformiert. Aus dieser Zelle kann das rekombinante Fusionsprotein isoliert werden. Die Durchführung kann beispielsweise wie in Beispiel 8 beschrieben erfolgen.

40 35 Diese Antikörper können beispielsweise zur Aufreinigung der entsprechenden Semaphorine, z.B. von H-SemaL und dessen Derivaten z.B. über Affinitätssäulen oder zum immunologischen Nachweis der Proteine, z.B. im ELISA, im Western-Blot und/oder in der Immunhistochemie verwendet werden. Die Antikörper können auch zur Analyse der Expression von H-SemaL, z. B. in verschiedenen Zelltypen bzw. Zelllinien verwendet werden.

45 40 Die cDNA von H-SemaL hat eine Länge von 2636 Nukleotiden (Tabelle 2). Das Genprodukt der H-SemaL-cDNA hat eine Länge von etwa 666 Aminosäuren (Tabelle 4) und weist die typische Domänenstruktur eines Semaphorins vom Typ L auf. Das Genprodukt weist ein N-terminales Signalpeptid (Aminosäuren 1 bis 44), Sema-Domäne (Aminosäure 45 bis ungefähr Aminosäure 545) und Ig (Immunglobulin)-Domäne (etwa Aminosäuren 550 bis 620) sowie am C-terminalen Ende eine hydrophobe Aminosäuresequenz auf, die eine potentielle Transmembrandomäne darstellt. Diese Domänen-Struktur wurde bisher für Semaphorine noch nie beschrieben. Es handelt sich um ein membranassoziiertes, wahrscheinlich an der Zelloberfläche lokalisiertes Glykoprotein einer neuen Untergruppe. Aufgrund dieser bisher nicht bekannten Domänenstruktur können die Semaphorine nun in VI Untergruppen eingeteilt werden:

I	Sezerniert, ohne weitere Domäne (z.B. ORF-A49)
II	Ig Sezerniert (ohne Transmembrandomäne) (z.B. AHV-Sema)
50 III	Ig, TM, CP Membranverankert mit zytoplasmatischer Sequenz (z.B. CD100)
IV	Ig, (P), HPC Sezerniert mit hydrophilem C-Terminus (z.B. H-Sema III, M-SemaD, Collapsin-1)
V	Ig, TM, CP Membranverankert mit C-terminalem 7 Thrombospondin-Motiv (z.B. M-Sema-F und G)
VI	Ig, TM Membranverankert (z.B. H-SemaL, M-SemaL)

55 55 Die nichtglykosyierte, nichtprozessierte Form von H-SemaL hat ein errechnetes Molekulargewicht von etwa 74,8 kD (74823 Dalton) (berechnet mit PeptideSort, GCG-Programm-Paket). Der Isoelektrische Punkt berechnet sich zu pH= 7,56.

Eine mögliche Signalpeptid-Schnittstelle liegt zwischen den Aminosäuren 44 und 45 (Tabelle 3; berechnet mit SignalP

(http://www.cbs.dtu.dk/services/Signal_P), einem auf neuronalen Netzwerken basierenden Programm zur Analyse von Signalsequenzen (Nielsen H. et al. (1997) Protein Engineering 10:1-6)). Dies ergibt für das prozessierte Protein (ohne Signalpeptid) ein Molekulargewicht (MW) von 70,3 kD (70323 Dalton) und einen Isoelektrischen Punkt von pH=7.01.

Die genomische Struktur ist ebenfalls weitgehend geklärt. Das H-SemaL-Gen weist 13 oder 15 oder mehr Exons, vorzugsweise 14 Exons und 12 oder 14 Introns, vorzugsweise 13 Introns auf. Aufgrund dieser komplexen Exon-Intron-Struktur sind unterschiedliche Splice-Varianten möglich. Die mRNA des transkribierten H-SemaL-Gens findet sich im Northern-Blot vor allem in Placenta, Keimdrüsen, Thymus und Milz. Es wurde keine mRNA in neuronalem Gewebe oder in Muskelgewebe nachgewiesen. Ein Hinweis auf eine spezifisch regulierte Expression in Endothelzellen liegt vor.

Durch alternatives Splicing können auch Formen von H-SemaL mit intrazytoplasmatischen Sequenzen entstehen, die eine Rolle in der intrazellulären Signaltransduktion spielen, ähnlich wie z.B. bei CD100.

Ebenfalls möglich wären durch alternatives Splicing entstehende, sezernierte Formen von H-SemaL, analog zum viralen AHV-Sema.

Nukleotid- und Aminosäuresequenzanalysen wurden mit Hilfe des GCG Programm-Pakets (Genetics Computer Group (1991) Program manual for the GCG package, version 7, 575 Science Drive, Wisconsin, USA 53711), FASTA (Pearson und Lipman (1988) Proc. Natl. Acad. Sci. 85, 2444-2448) und BLAST-Program (Gish und States (1993) Nat. Genet. 3, 266-272; Altschul et al. (1990) J. Mol. Biol. 215, 403-410) durchgeführt. Diese Programme wurden auch für Sequenzvergleiche mit GenBank (Version 102.0) und Swiss Prof (Version 34.0) verwendet.

Posttranskriptionale Modifikationen wie Glykosylierung und Myristylierung von H-SemaL sind ebenfalls möglich. Konsensus-Sequenzen für N-Glykosylierungsstellen wurden mit Hilfe des Programms Prosite (GCG Programm-Paket) an den Positionen 105, 157, 258, 330, 602 der Aminosäuresequenz von H-SemaL (gemäß Tabelle 4) gefunden, solche für Myristylierung an den Positionen 114, 139, 271, 498, 499, 502, 654 (Konsensus-Sequenz: G~(E, D, R, K, H, P, F, Y, W) x (S, T, A, G, C, N)~(P)). Darüber hinaus enthält die Aminosäuresequenz von H-SemaL mehrere Konsensus-Sequenzen für potentielle Phosphorylierungsstellen durch unterschiedliche Kinasen. Deshalb kann davon ausgegangen werden, daß H-SemaL das Substrat unterschiedlicher Kinasen sein kann, z.B. Phosphorylierungsstellen für Kreatin-Kinase 2, Protein-Kinase C und Tyrosin-Kinase.

Vorausgesagte Kreatin-Kinase 2-Phosphorylierungs-Stellen (Konsensussequenz CK2: (S,T)x2(D,E)) (Prosite, GCG) an den Positionen 119, 131, 173, 338, 419, 481 der Aminosäuresequenz. Vorausgesagte Protein-Kinase-C-Phosphorylierungs-Stellen (Konsensussequenz PkC: (S,T)x(R,K)) (Prosite, GCG) an den Positionen 107, 115, 190, 296, 350, 431, 524, 576 der Aminosäuresequenz. Vorausgesagte Tyrosin-Kinase-Phosphorylierungs-Stelle (Konsensussequenz: (R,K)x(2,3)(D,E)x(2,3)Y) (Prosite GCG) an Position 205 der Aminosäuresequenz.

Die Konsensussequenzen sind im Einbuchstabencode für Aminosäuren angegeben.

Ein für Integrine charakteristisches "RGD"-Motiv (Arginin-Glycin-Asparaginsäure) findet sich an Position 267.

Die Glykosylierungsstellen sind gut konserviert zwischen viralem AHV-Sema, H-SemaL und (soweit bekannt) M-SemaL.

Eine Di- oder Multimerisierung von H-SemaL ist möglich und wurde bei anderen Semaphorinen wie CD100 beschrieben (Hall et al. (1996)). Das CD100 Molekül ist ebenfalls ein membranverankertes Glykoprotein-Dimer von 150kd. CD100 ist jedoch nicht nahe verwandt mit dem erfundungsgemäßen humanen Semaphorin (H-SemaL).

Die partielle cDNA-Sequenz von M-SemaL hat eine Länge von 1195 Nukleotiden. Diese Sequenz kodiert für ein Protein mit 394 Aminosäuren. Diese 394 Aminosäuren entsprechen den Aminosäuren 1 bis 396 von H-SemaL. Das Signalpeptid im M-SemaL ersteckt sich über die Aminosäuren 1 bis 44 (genau wie im H-SemaL). Die Sema-Domäne beginnt bei der Aminosäure 45 und erstreckt sich bis zum Ende bzw. wahrscheinlich über das Ende der Sequenz gemäß Tabelle 4 hinaus.

Multiple Alignments wurden mit Hilfe des Programms Clustal W (Thompson et al. (1994)) durchgeführt. Diese Alignments wurden manuell weiterbearbeitet mit Hilfe von SEAVIEW (Galtier et al. (1996) Comput Appl. Biosci 12, 543-548). Die phylogenetischen Entferungen wurden mit Clustal W (Thompson et al. (1994)) bestimmt.

Ein Vergleich der Proteinsequenzen der bekannten und der neuen Semaphorine und eine phylogenetische Analyse dieser Sequenzen zeigt, daß sich die Gene entsprechend ihrer phylogenetischen Verwandtschaft einteilen lassen. Hier fließt natürlich die C-terminale Domänenstruktur der entsprechenden Semaphorin-Subtypen als entscheidender Faktor mit ein, weshalb Semaphorine der gleichen Untergruppen in der Regel auch phylogenetisch näher verwandt sind, als Semaphorine unterschiedlicher Untergruppen. Einfluß hat auch, aus welcher Spezies das Semaphorin isoliert wurde, d.h. ob die entsprechenden Spezies phylogenetisch nahe miteinander verwandt sind oder nicht.

Eine phylogenetische Analyse (vergl. Figur 3) der bekannten Semaphorin Aminosäuresequenzen (vollständige Sequenzen und/oder Teilsequenzen, wobei die Aminosäuresequenzen für H-SemaL und M-SemaL gemäß den Tabellen 4 und 5 verwendet wurden, für alle anderen Sequenzen, die unter den Zugangsnummern gespeicherten Sequenzen bzw. die von diesen Sequenzen abgeleiteten Aminosäuresequenzen) mittels des Programms CLUSTALW (Thompson J.D. et al. (1994) Nucleic Acids Res. 22:4673-4680) zeigt, daß die Aminosäuresequenzen von H-SemaL und M-SemaL phylogenetisch nahe miteinander verwandt sind und eine eigene phylogenetische Gruppe bilden. H-SemaL und M-SemaL wiederum sind phylogenetisch am nächsten verwandt mit AHV-Sema und Vac-A39. Sie sind

unteneinander deutlich näher verwandt, als mit irgendeinem anderen bisher bekannten Semaphorin. Die Analyse zeigt weiterhin, daß auch andere Semaphorine phylogenetisch nahe miteinander verwandt sind und eigene Gruppen innerhalb der Semaphorine bilden. Beispielsweise fallen die Semaphorine, die sezerniert werden, z.B. H-Sema III, IV, V, und E in eine phylogenetische Gruppe. Zu dieser Subfamilie gehören auch deren Homologe in anderen Spezies, während das humane (transmembrane) CD100 mit dem entsprechenden Maus-Homologen (M-Sema G2) und mit Collapsin-4 in eine phylogenetische Gruppe fällt.

5 Im Bezug auf die gesamten Aminosäuresequenzen liegen die beobachteten Homologien innerhalb der phylogenetischen Gruppen zwischen etwa 90% und 80% Aminosäureidentität im Bezug auf sehr nahe verwandte Gene wie z.B. H- und M-SemaE oder -III/D und etwas weniger als 40% bei wenig verwandten Genen der Semaphorine. Innerhalb der 10 Sema-Domäne liegt die beobachtete Aminosäureidentität um einige Prozent höher, und durch ihren großen Anteil am Gesamtprotein (50-80% des Proteins gehören zur Sema-Domäne) der Aminosäuresequenz beeinflußt diese wesentlich die Gesamtidentität.

15 H-SemaL ist, über das Gesamtprotein berechnet, zu 46% identisch mit AHV-Sema, wird dagegen die Sema-Domäne allein betrachtet, dann beträgt die Aminosäureidentität 53%. Dies ist höher als z.B. zwischen den verwandten M-SemaB und -C (37% Identität im Bezug auf das Gesamtprotein, 43% Identität im Bezug auf die Sema-Domäne), ähnlich wie M-SemaA und -E (43% Gesamtprotein, 53% Sema-Domäne). Die Aminosäureidentität zwischen der partiellen 20 M-SemaL Sequenz (Tabelle 6) und H-SemaL (Tabelle 5) liegt im Bereich der Sema-Domäne bei 93%, so daß davon ausgegangen werden kann, daß es sich um das entsprechend homologe Gen der Maus handelt.

Korrespondierende Semaphorine zu H-SemaL und M-SemaL in anderen Spezies können innerhalb der Sema-Domäne 25 eine Aminosäureidentität größer als 40% im Bezug auf H-SemaL aufweisen. Bei den nahe verwandten Wirbeltieren (Säuger, Vögel) können sogar Aminosäureidentitäten über 70% angetroffen werden.

Es handelt sich um eine neue Subfamilie von Semaphorinen mit größerer Aminosäureidentität zu dem viralen AHV-Sema als zu den bisher bekannten humanen bzw. murinen Semaphorinen, und mit einer für humane Semaphorine bisher nicht bekannten C-terminalen Struktur. Diese neuen Semaphorine (Mitglieder der Subfamilie) zeichnen sich 30 dadurch aus, daß sie aufgrund ihrer Domänenstruktur in die Untergruppe IV fallen und/oder die gleiche phylogenetische Gruppe fallen wie H-SemaL und M-SemaL und/oder im Bezug auf die gesamte Aminosäuresequenz zu H-SemaL eine Aminosäureidentität von mindestens 30 bis 40 %, vorzugsweise 50 bis 60 %, besonders bevorzugt 70 bis 80 % oder eine größere Identität aufweisen und/oder im Bezug auf die Sema-Domäne eine Aminosäureidentität mit H-SemaL von mindestens 70 %, vorzugsweise größer 80 %, besonders bevorzugt größer 90 % aufweisen.

35 30 Den Semaphorinen vom Typ L kommt auch eine andersartige biochemische Funktion zu. Eine neue Funktion dieser Semaphorine liegt in der Modulation des Immunsystems.

Das nächste Verwandte von H-SemaL ist das virale AHV-Semaphorin (AHV-Sema). Dieses ist von ähnlicher Größe, besitzt aber im Gegensatz zum H-SemaL keine Transmembrandomäne. Das AHV-Sema wird vermutlich von 40 virusinfizierten Zellen sezerniert, um den H-SemaL äquivalenten Rezeptor (Semaphorin von Typ L im Streifengnu) im natürlichen Wirt (Streifengnu) zu blockieren, und so dem Angriff des Immunsystems zu entgehen. Ferner ist eine Funktion als repulsives Agens (Chemorepellent) für Zellen des Immunsystems denkbar.

45 Die biochemische Funktion der neuen Semaphorine vom Typ L und deren Derivate ist als generell immunmodulierend und/oder entzündungsmodulierend anzusehen. Einerseits können sie

50 40 A) als die Immunantwort hemmende Moleküle entweder lokal, z.B. als Transmembranprotein an der Oberfläche von Zellen oder auch über größere Entfernung, z.B. wenn sie durch Prozessierung (z.B. Proteasen) oder alternatives Splicing sezerniert werden, z.B. durch Diffusion im Gewebe, ihre Wirkung als Chemorepellent und/oder Immunsuppressivum entfalten.

55 45 Beispielsweise kann die Expression dieser neuen Semaphorine vom Typ L z.B. an der Oberfläche der Zellen der Gefäßendothelien das Leukozyten-Attachment und deren Migration durch die Gefäßwand verhindern. Den neuen Semaphorinen kann eine Rolle bei der Aufrechterhaltung von Schrankenwirkungen, z.B. zur Verhinderung von Infektionen in besonders "wichtigen" oder exponierten Organen, beispielsweise zur Aufrechterhaltung der Blut-Hirn-Schranke, des Plazentarkreislaufs und/oder anderen immunologisch privilegierten Orten (z.B. Pancreas-Inseln) und/oder beim Schutz vor Autoimmunerkrankungen zukommen. Darüber hinaus können die neuen Semaphorine und/oder ihre Derivate in verschiedenen Geweben auch an repulsiven Signalen, beispielsweise für Zellen des Immunsystems (z.B. Leukozyten) als Schutz gegen versehentliche Aktivierung von Abwehrmechanismen beteiligt sein.

60 50 B) Weiterhin können den neuen Semaphorinen und/oder deren Derivaten Funktionen als akzessorische Moleküle zukommen. An der Zelloberfläche exprimiert können sie beispielsweise an der Interaktion mit Zellen des Immunsystems im Rahmen der Aktivierung von Abwehrmechanismen z. B. bei Virusinfektionen beteiligt sein.

65 55 Dadurch ergeben sich mehrere Verwendungsmöglichkeiten für die neuen Semaphorine vom Typ L und deren Deri-

vaten sowie den für diese Proteine kodierenden Nukleinsäuren.

Funktion A): Es handelt sich um ein immunsuppressives und/oder entzündungshemmendes Prinzip: Zahlreiche potentielle Anwendungsmöglichkeiten liegen in den Bereichen Organtransplantation, Entzündungstherapie, Immuntherapie und Gentherapie.

5 Beispielsweise können mit Hilfe der Semaphorin-kodierenden DNA oder Derivaten derselben nichthumane, transgene Tiere hergestellt werden.

Eine Anwendungsmöglichkeit dieser Tiere liegt in der Hemmung der Transplantatabstoßung in transgenen Modellen für Organtransplantationen. Beispielsweise können transgene, gegen Abstoßung geschützte tierische Organe für Xenotransplantationen hergestellt werden. Dies sollte z.B. auch zusammen mit anderen Transgenen (z.B. Komplementregulatoren wie DAF oder CD59) möglich sein.

10 Eine weitere Anwendung liegt in der Herstellung von nicht-humanen Knock-out Tieren, beispielsweise von Knock-out Mäusen („Laboratory Protocols for Gene-Targeting“, Torres and Kühn (1997) Oxford University Press, ISBN 0-19-963677-X): Durch Knock-out des Mausgens M-SemaL können z.B. weitere Funktionen des Gens aufgefunden werden: Sie stellen auch potentielle Modellsysteme für entzündliche Erkrankungen dar, falls die Mäuse ohne Semaphorin-Gen

15 lebensfähig sind. Sollte M-SemaL für die Immunmodulation wichtig sein, so sind vermehrt solche Mäuse zu erwarten. Weiterhin können nicht-humanen Knock-in-Tiere, beispielsweise Mäuse, hergestellt werden. Dabei wird z.B. M-SemaL durch normales/verändertes H-SemaL oder verändertes M-SemaL (z.B. Integration der neuen Semaphorin-Subtypen unter der Kontrolle von konstitutiven und/oder induzierbaren Promotoren) ersetzt. Solche Tiere können z.B. für die Suche nach weiteren Funktionen der neuen Semaphorine, z.B. Funktionen des humanen Gens oder Derivaten dieser

20 Gene dienen oder zur Identifizierung und Charakterisierung von immunmodulierenden Wirkstoffen benutzt werden.

25 Verwendung von z. B. Nukleinsäuren, die für Semaphorine vom Typ L oder Derivate derselben kodieren, zur Herstellung von z.B. rekombinanteren Immunsuppressiven, anderen löslichen Proteinen oder Peptiden die sich von der Aminosäuresequenz der Semaphorine vom Typ L, z.B. Von H-SemaL oder den entsprechenden Nukleinsäuren, z.B. Genen ableiten. In ähnlicher Weise können auch Agonisten mit struktureller Ähnlichkeit hergestellt werden. Diese immunsuppressiven Wirkstoffe/Agonisten können bei Autoimmunerkrankungen und entzündlichen Erkrankungen und/oder Organtransplantationen eingesetzt werden.

30 Gentherapie mit Semaphorinen vom Typ L, z.B. mit Nukleinsäuren, die für H-SemaL oder deren Derivate kodieren, z. B. mittels viraler oder nichtviralier Methoden. Einsatz bei Autoimmunerkrankungen und entzündlichen Erkrankungen, der Transduktion von Organen sowie vor/während/nach Transplantationen zur Verhinderung der Transplantatabstoßung.

35 Insbesondere können die neuen Semaphorine und/oder die für diese Semaphorine kodierenden Nukleinsäuren und Derivate derselben, insbesondere H-SemaL, für H-SemaL kodierende DNA und Derivate derselben in einem Verfahren zum Screening von Wirkstoffen, insbesondere zur Identifizierung und Charakterisierung von immunmodulierenden Wirkstoffen, eingesetzt werden.

40 Funktion B): H-SemaL ist ein akzessorisches Molekül, das an der Zelloberfläche exprimiert wird und an der Interaktion mit Zellen, z.B. des Immunsystems, z.B. als akzessorisches Molekül in der Aktivierung von Signalwegen, beteiligt ist. Ein virales Gen bzw. das Genprodukt eines viralen oder anderen pathogenen Gens z.B. mikrobiologischen Ursprungs könnte z.B. als kompetitiver Inhibitor dieses akzessorischen Moleküls wirken. Eine Anwendung für die neuen Semaphorine liegt bei dieser Funktion ebenfalls im Bereich der Organtransplantation, Entzündungstherapie, Immuntherapie und/oder Gentherapie.

45 Beispielsweise können die neuen Semaphorine in einem Verfahren zum Screening von antagonistischen Wirkstoffen bzw. Inhibitoren verwendet werden. Auf diese Weise identifizierte Wirkstoffe können dann z.B. zur Blockade des Semaphorin-Rezeptors eingesetzt werden. Lösliche und/oder sezernierte H-SemaL Antagonisten bzw. Inhibitoren können beispielsweise chemische Substanzen oder die neuen Semaphorine bzw. Derivate derselben selbst sein (z.B.

50 Teile/verkürzte Formen derselben, beispielsweise ohne Membrandomäne oder als Ig-Fusionsproteine oder von diesen abgeleitete Peptide, die geeignet sind, den korrespondierenden Rezeptor zu blockieren). Auf diese Weise identifizierte, spezifische Antagonisten und/oder Inhibitoren können beispielsweise kompetitiv wirken und bei der Hemmung der Abstoßung z.B. in transgenen Modellen für Organtransplantationen und bei Autoimmunerkrankungen, entzündlichen Erkrankungen und Organtransplantationen eingesetzt werden. Nukleinsäuren, z. B. DNA, die für die neuen Semaphorine kodieren bzw. deren mit Hilfe von molekulärbiologischen Methoden erzeugte Derivate können beispielsweise für die Herstellung nichthumane, transgener Tiere verwendet werden. Eine Überexpression von H-SemaL kann in diesen transgenen Tieren zu vermehrter Anfälligkeit für Autoimmunerkrankungen und/oder entzündlichen Erkrankungen führen. Solche transgenen Tiere eignen sich damit zum Screening von neuen, spezifischen, immunmodulierenden Wirkstoffen.

55 Solche Nukleinsäuren können ebenso für die Herstellung von nicht-humanen Knock-out Tieren, beispielsweise Knock-out Mäusen, bei denen das Mausgen M-SemaL ausgeschaltet wird, verwendet werden. Solche Knock-out Tiere können für die Suche nach weiteren biochemischen Funktionen des Gens eingesetzt werden. Sie stellen auch potentielle Modellsysteme für entzündliche Erkrankungen dar, falls die Mäuse ohne das M-SemaL Gen lebensfähig sind.

Diese DNA kann ebenso zur Herstellung von nicht-humanen Knock-in Tieren, beispielsweise Mäusen verwendet werden. Dabei wird das M-SemaL-Gen durch ein verändertes M-SemaL Gen/cDNA oder ein gegebenenfalls verändertes, z.B. mutiertes Semaphorin Typ L Gen/cDNA einer anderen Spezies, z.B. von H-SemaL ersetzt. Solche transgenen Tiere können für die Suche nach weiteren Funktionen der erfundungsgemäßen Semaphorine verwendet werden.

5 Die Erfindung betrifft auch die Verwendung der Semaphorine vom Typ L und deren Derivate, sowie der für diese Proteine kodierenden Nukleinsäuren, z. B. Gene/cDNAs und deren Derivate und/oder mit Hilfe dieser Semaphorine identifizierter Wirkstoffe zur Herstellung von Arzneimitteln. Beispielsweise können Arzneimittel hergestellt werden, die in der Gentherapie angewendet werden können und die Agonisten und/oder Antagonisten der Expression der Semaphorine vom Typ L, beispielsweise von H-SemaL, enthalten. Dazu können z. B. virale und/oder nichtvirale Methoden 10 verwendet werden. Diese Arzneimittel können z.B. bei Autoimmunerkrankungen und entzündliche Erkrankungen, Organtransplantationen vor und/oder während und/oder nach der Transplantation zur Verhinderung der Abstoßung eingesetzt werden.

10 Die für die neuen Semaphorine kodierenden Nukleinsäuren, z. B. Gene, cDNAs und deren Derivate können auch als Hilfsmittel in der Molekularbiologie eingesetzt werden.

15 Darüberhinaus können die neuen Semaphorine, insbesondere H-SemaL und Nukleinsäuren, z. B. Gene/cDNAs derselben in Verfahren zum Screening neuer Wirkstoffe eingesetzt werden. Modifizierte Proteine und/oder Peptide, die sich z. B. von H-SemaL und/oder M-SemaL ableiten, können zur Suche nach dem entsprechenden Rezeptor und/oder dessen Antagonisten bzw. Agonist in funktionellen Tests, beispielsweise mittels Expressionskonstrukten von H-SemaL und Homologen eingesetzt werden.

20 Die Erfindung betrifft auch die Verwendung eines Semaphorins vom Typ L oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert in einem Verfahren zur Identifizierung von pharmakologischen Wirkstoffen, insbesondere von immunmodulierenden Wirkstoffen.

Die Erfindung betrifft auch Verfahren zur Identifizierung von Wirkstoffen, wobei ein Semaphorin vom Typ L oder ein Derivat derselben bzw. eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder ein Derivat derselben eingesetzt wird, um pharmakologische Wirkstoffe, z. B. immunmodulierende Wirkstoffe zu identifizieren. Beispielsweise betrifft die Erfindung ein Verfahren, bei dem ein Semaphorin vom Typ L unter definierten Bedingungen mit einem zu untersuchenden Wirkstoff inkubiert wird und parallel ein zweiter Ansatz ohne den zu untersuchenden Wirkstoff, aber unter ansonsten gleichen Bedingungen durchgeführt wird und dann die inhibierende bzw. aktivierende Wirkung des zu untersuchenden Wirkstoffs bestimmt wird.

25 Beispielsweise betrifft die Erfindung auch Verfahren zur Identifizierung von Wirkstoffen, wobei eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder ein Derivat derselben unter definierten Bedingungen in Gegenwart eines zu untersuchenden Wirkstoffs exprimiert wird und das Ausmaß der Expression bestimmt wird. Gegebenenfalls können auch in einem solchen Verfahren zwei oder mehrere Ansätze parallel unter gleichen Bedingungen durchgeführt werden, wobei die Ansätze aber unterschiedliche Mengen des zu untersuchenden Wirkstoffs enthalten.

30 Beispielsweise kann der zu untersuchende Wirkstoff die Transkription und/oder die Translation inhibieren oder aktivieren.

Das Semaphorin vom Typ L kann wie seine viralen Homologen an das neu beschriebene Rezeptormolekül VESPR (Comeau et al, (1998) *Immunity*, Vol. 8, 473-482) binden und kann vermutlich in Monozyten eine Induktion von Zell-Adhäsionsmolekülen wie ICAM-1 und Zytokinen wie Interleukin-6 und Interleukin-8 bewirken. Dies kann zu deren Aktivierung und zur Zellaggregation führen. Das Expressionsmuster des VESPR-Rezeptors zeigt teilweise interessante Parallelen zu H-Senal, z.B. eine starke Expression in Placenta und eine deutliche Expression in Milzgewebe. Interaktionen mit weiteren, noch unbekannten Rezeptoren der Plexin-Familie, oder anderen Rezeptoren sind möglich. Auch eine Interaktion mit sich selbst oder anderen semaphorinähnlichen Molekülen ist möglich. Eine Interaktion der Semaphorine vom Typ L kann insbesondere über eine konservierte Domäne im C-terminalen Bereich der Sema-Domäne stattfinden.

35

40

45 Zu den Annotation Plasmiden:

pMelBacA-H-SemaL (6622bp) in pMelBacA (Invitrogen, De Schelp, NL)(SEQ ID NO.42). Nukleotid 96-98 ATG - Startkodon, Nukleotid 96-168 Mellitin Signal-Sequenz, Nukleotid 168-173 BamHI Schnittstelle (PCR/Klonierung), Nukleotid 171-1998 Leserahmen SemaL Aminosäuren 42-649 (ohne eigene Signal-Sequenz und ohne Transmembransequenz), Nukleotid 1993-1998 EcoRI Schnittstelle (PCR/Klonierung) und Nukleotid 1992-1994 Stop Codon

50

55 Plasmid pCDNA3.1-H-SemaL-MychisA (7475 bp) (SEQ ID NO. 35):
Nukleotid 954-959 BamHI Schnittstelle (Klonierung),
Nukleotid 968-970 ATG SemaL, Nukleotid 968-2965 Leserahmen SemaL, Nukleotid 2963-2968 Pml I Schnittstelle, Nukleotid 2969-2974 HindIII Schnittstelle,
Nukleotid 2981-3013 Myc-Tag, Nukleotid 3026-3033 6xHis-Tag, Nukleotid 3034-3036 Stop Codon,

EP 0 892 047 A2

Plasmid pCDNA3.1-H-SemaL-EGFP-MychisA (8192 bp): (SEQ ID NO. 36):

Nukleotid 954-959 BamHI Schnittstelle (Klonierung), Nukleotid 968-970 ATG SemaL, Nukleotid 968-2965 Leserahmen SemaL, Nukleotid 2963-2965 halbe Pml I Schnittstelle, Nukleotid 2966-3682 Leserahmen EGFP (in Pml I Kloniert), Nukleotid 3683-3685 halbe Pml I Schnittstelle, Nukleotid 3685-3691 HindIII, Nukleotid 3698-3730 Myc-Tag, Nukleotid 3743-3760 6xHis-Tag, und Nukleotid 3761-3763 Stop Codon

5 Plasmid pIND-H-SemaL-EA (7108 bp) in Vektor pIND (Invitrogen, De Schelp, NL) (SEQ ID No. 38):

10 Nukleotid 533-538 BamHI Schnittstelle (Klonierung), Nukleotid 546-548 ATG SemaL, Nukleotid 546-Leserahmen SEMAL, Nukleotid 2542-2547 Pml I Schnittstelle, Nukleotid 2546-2553 HindIII Schnittstelle und Nukleotid 2563-2565 Stop Codon.

15 Plasmid pIND-H-SemaL-EE (Gesamtlänge 7102 bp) in Vektor pIND (Invitrogen, De Schelp, NL) (SEQ ID No. 37):

Nukleotid 533-538 BamHI Schnittstelle (Klonierung), Nukleotid 546-548 ATG SemaL, Nukleotid 546-Leserahmen SemaL, Nukleotid 2542-2547 Pml I Schnittstelle, Nukleotid 2548-2553 HindIII Schnittstelle, Nukleotid 2560-2592 Myc-Tag, Nukleotid 2605-2622 6xHis-Tag und Nukleotid 2623-2625 Stop Codon.

20 Plasmid pQE30-H-SemaL-179-378.seq (4019 bp) in Vektor pQE30 (Qiagen, Hilden) entspricht pQE30-H-SemaLBH (SEQ ID No. 39):

Nukleotid 115-117 ATG, Nukleotid 127-144 6xHis-Tag, Nukleotid 145-750 BamHI-HindIII PCR-Fragment SemaL Aminosäuren (aa) 179-378 und Nukleotid 758-760 Stop Codon.

25 Plasmid pQE31-H-SemaL- (SH (3999 bp) in Vektor pQE31 (Qiagen, Hilden) (SEQ ID No. 40):

Nukleotid 115-117 ATG, Nukleotid 127-144 6xHis-Tag, Nukleotid 147-152 BamHI, Nukleotid 159-729 SacI-HindIII Fragment SemaL (C-terminal) aa480-666 und Nukleotid 734-736 Stop Codon.

Beispiele:

Versuchsbedingungen, die in den Beispielen Anwendung finden:

30 Verwendete PCR-Programme:

Taq52-60 (mit Ampli-Taq^R-Polymerase, Perkin Elmer, Weil der Stadt, Deutschland)

35 96°C/60s 1 Zyklus
96°C/15s-52°C/20s-70°C/60s 40 Zyklen
70°C/60s 1 Zyklus

Taq60-30

40 96°C/60s 1 Zyklus
96°C/15s-60°C/20s-70°C/30s 35 Zyklen
70°C/60s 1 Zyklus

Taq60-60

45 96°C/60s 1 Zyklus
96°C/15s-60°C/20s-70°C/60s 35 Zyklen
70°C/60s 1 Zyklus

50 Taq62-40

96°C/60s 1 Zyklus
96°C/15s-62°C/20s-70°C/40s 35 Zyklen
70°C/60s 1 Zyklus

55 Verwendete Reaktionsbedingungen für PCR mit Taq-Polymerase:

50µl Reaktionsansätze mit 100-200ng Template, 200µM dNTP, 0,2-0,4 µM je Primer, 2,5U Ampli-Taq^R, 5µl des mit-

gelieferten 10x Reaktionspuffers

Verwendete Programme für:

5 1. XL62-6 (mit Expand-Long Template PCR System^R, Boehringer Mannheim, Deutschland)

94°C/60s	1 Zyklus
94°C/15s-62°C/30s-68°C/6min	10 Zyklen
94°C/15s-62°C/30s-68°C/(6min+15s/Zyklus)	25 Zyklen
10 68°C/7min	1 Zyklus

2. XL62-12 (mit Expand-Long Template PCR System^R, Boehringer Mannheim, Deutschland)

15 94°C/60s	1 Zyklus
94°C/15s-62°C/30s-68°C/12min	10 Zyklen
94°C/15s-62°C/30s-68°C/(12min+15s/Zyklus)	25 Zyklen
68°C/7min	1 Zyklus

Reaktionsbedingungen für PCR mit Expand-Long Template PCR System

20 50µl Reaktionsansätze mit 100-200ng Template, 500µM dNTP, 0,2-0,4 µM je Primer, 0,75µl Enzym-Mix, 5µl des mitgelieferten 10x Reaktionspuffers Nr. 2.

Beispiel 1:

25 Ausgehend von Sequenzen des AHV-Sema (Ensser u. Fleckenstein (1995), J. General Virol. 76: 1063-1067) wurden den PCRs und RACE-PCRs durchgeführt. Als Ausgangsmaterial hierfür diente humane cDNA aus Plazenta-Gewebe, an welche Adapter zur RACE-Amplifikation ligiert wurden (Marathon™-cDNA Amplification Kit, Clontech Laboratories GmbH, Tullastraße 4, 69126 Heidelberg, Deutschland). Zunächst wurde mittels spezifischer Primer (Nr. 121234 + Nr. 30 121236, Tabelle 6) ein PCR-Fragment mit einer Länge von etwa 800bp (Basenpaaren) amplifiziert (PCR-Programm: (Taq60-60)). Dieses wurde kloniert und sequenziert (Taq-Dye-Deoxy-Terminator Sequenzierungs-Kit, Applied Biosystems, Foster City, CA USA) Brunnenweg 13, Weilderstadt). Die Sequenzierung des PCR-Produkts ergab eine Sequenz, die eine hohe Homologie zu der DNA-Sequenz von AHV-Sema aufweist, identisch zu der Sequenz der bei- den ESTs.

35 Ein PCR-Fragment von 600bp wurde mit dem Primerpaar (Nr. 121237 + Nr. 121239, Tabelle 6) identifiziert. Es zeigte sich, das es sich um Klone mit DNA-Sequenzen des selben Gens handelte.

Beispiel 2:

40 Das 800bp PCR-Fragment aus Beispiel 1 wurde radioaktiv markiert (Random-Priming nach der Methode von Feinberg (1983) Anal. Biochem. 132:6-13), mit ³²P- α -dCTP) und als Sonde für einen Multi-Tissue Northern-Blot (Human Multiple Tissue Northern Blot II, Clontech, Heidelberg, Germany), der mRNA-Proben aus den Geweben Milz, Thymus, Prostata, Testis, Ovarien, Dünndarm, Dickdarm und Leukozyten (PBL) enthält, verwendet. Dabei zeigte sich deutlich die Expression einer mRNA mit einer Länge von etwa 3.3kb in Milz und Keimdrüsen (Hoden, Eierstöcken), 45 sowie schwächer in Thymus und Darm. Eine Hybridisierung eines Master-Blots (Dotblot mit RNA aus zahlreichen Geweben (Human RNA Master Blot™, Clontech) bestätigte dieses Ergebnis und zeigte auch eine starke Expression in Plazenta-Gewebe.

Die Hybridisierung wurde für 16 Stunden unter stringenten Bedingungen (5xSSC, 50 mM Na-Phosphat pH 6.8, 50 % Formamid, 100 µg/ml Hefe-RNA) bei 42°C durchgeführt. Die Blots wurden stringent gewaschen (65°C, 0,2XSSC, 0.1 % SDS) und einem Fuji BAS2000 Phosphoimager™ exponiert

Beispiel 3:

55 Eine cDNA-Bibliothek aus humaner Milz, kloniert in dem Bakteriophagen Lambda gt10 (Human Spleen 5' STRETCH PLUS cDNA, Clontech) wurde mit dieser Sonde durchsucht und ein Lambda-Klon identifiziert. Die in diesem Klon inserierte cDNA mit einer Länge von 1.6kb wurde mittels PCR (Expand™ Long Template PCR System, Boehringer Mannheim GmbH, Sandhofer Straße 116, 68305 Mannheim) amplifiziert wobei die vektorspezifischen Primer Nr. 207608 + Nr. 207609 (Tabelle 6) verwendet wurden (flankierend der EcoRI-Klonierungsstelle) und das erhaltene PCR-

Fragment sequenziert. Dieser Klon enthielt das 5' Ende der cDNA und erweiterte die bekannte cDNA Sequenz auch nach 3'. Ausgehend von den neuen Teilsequenzen der cDNA wurden neue Primer für die RACE-PCR entwickelt (Nr. 232643, Nr. 232644, Nr. 233084, Tabelle 6). Zusammen mit einer verbesserten Thermocyclertechnik (PTC-200 von MJ-Research, Biozym Diagnostik GmbH, 31833 Hess. Oldendorf) mit deutlich besseren Leistungsdaten (Heiz- und Kühlrate) wurde ein 3'-RACE PCR-Produkt amplifiziert wobei die Primer Nr. 232644 bzw. Nr. 232643 und AP1 verwendet wurden und in den Vektor pCR2.1 (Invitrogen, De Schelp 12, 9351 NV Leek, Niederlande) kloniert. Das 3'-RACE PCR-Produkt wurde Sequenziert und auf diese Weise das 3' Ende der cDNA identifiziert. Eine RACE-Amplifikation nach 5' (Primer Nr. 131990 bzw. Nr. 233084 und AP1) erweiterte das 5' Ende der cDNA um wenige Nukleotide und bestätigte den im identifizierten Lambda-Klon gefundenen Aminoterminal des H-SemaL.

10

Beispiel 4:

Ausgehend von einem kurzen murinen EST (Zuordnungs-Nr. AA260340) und einem daraus abgeleiteten Primer Nr. 260813 (Tabelle 6) und dem H-SemaL spezifischen Primern Nr. 121234 (Tabelle 6) wurde mittels PCR (Bedingungen: Taq52-60) ein DNA-Fragment mit einer Länge von ca. 840 bp muriner cDNA amplifiziert und in den Vektor pCR2.1 kloniert. Das dieses DNA-Fragment enthaltende Gen wurde M-SemaL genannt. Mit dem erhaltenen M-SemaL DNA-Fragment eine cDNA-Bank aus Mäuse-Milz (Mouse Spleen 5' STRETCH cDNA, Clontech) untersucht, wobei bereits mehrere Klone identifiziert werden konnten.

15 PCR (Taq60-30) mit den Primern Nr. 260812, Nr. 260813 aus muriner, endothelialer cDNA lieferte ein PCR-Fragment mit einer Länge von 244 Basenpaaren. Die PCR-Ergebnisse zeigten, daß eine deutliche basale Expression in murinen Endothelzellen vorhanden ist, welche nach Stimulation mit dem Zytokin Interferon- γ und Lipopolysacchariden zurückgeht.

20

Beispiel 5:

25

Untersuchungen zur Chromosomalen Lokalisation wurden mittels Fluoreszenz in-situ Hybridisierung (FISH) durchgeführt. Dazu wurden Metaphase Chromosomen von Mensch und Maus, ausgehend von einer humanen Blutprobe bzw. der Mauszelllinie BINE 4.8 (Keyna et al. (1995) J. Immunol. 155, 5536-5542), hergestellt (Kraus et al. (1994) Genomics 23, 272-274). Die Objekträger wurden mit RNase und Pepsin behandelt (Liehr et al. (1995) Appl. Cytogenetics 21, 185-188). Für die Hybridisierung wurden 120 ng humane Nick-translatierte Semaphorin-Probe bzw. 200 ng einer entsprechenden Mausprobe verwendet. Die Hybridisierung wurde jeweils in Gegenwart von 4.0 μ g COT1-DNA und 20 μ g STD bei 37°C (3 Tage) in einer befeuchteten Kammer durchgeführt.

30 Die Objekträger wurden mit 50% Formamid/2x SSC (3 mal je 5 min bei 45°C) und dann mit 2 x SSC (3 mal je 5 min bei 37°C) gewaschen und die biotinylierte Probe mit dem FITC-Avidin-System (Liehr et al. (1995)), detektiert. Die Objekträger wurden mit Hilfe eines Fluoreszenz-Mikroskops ausgewertet. Es wurden 25 Metaphasen/Probe ausgewertet, wobei jedes Experiment doppelt durchgeführt wurde. Es zeigt sich, daß H-SemaL auf Chromosom 15q23 lokalisiert ist Chromosomal benachbart liegen der Locus für das Bardet-Biedl-Syndrom und Tay-Sachs Erkrankung (Hexosaminidase A).

40

Beispiel 6:

Die genomische Intron-Exon Struktur des H-SemaL Gens ist zum größten Teil aufgeklärt. Genomische DNA Fragmente wurden ausgehend von 250 mg humaner genomischer DNA, die aus PHA stimulierten peripheren Lymphozyten (Blut) isoliert worden waren, amplifiziert. Kürzere Fragmente wurden mit Ampli Taq^R (Perkin Elmer), längere Fragmente mit dem Expanded Long Templat PCR System^R (Boehringer Mannheim) amplifiziert.

45 Durch PCR-Amplifikation konnte bisher fast der vollständige genomische Locus des H-SemaL kloniert und charakterisiert werden. Insgesamt konnten bereits mehr als 8888 bp der genomischen Sequenz bestimmt werden und so die Intron-Exon-Struktur des Gens weitgehend aufgeklärt werden.

50

Beispiel 7:

Expressionsklonierungen:
Da kein kompletter Klon des Semaphoringens aus der Lambda-gt10 cDNA-Bank isoliert werden konnte, und auch mittels PCR ein vollständiger Klon nicht zu erhalten war, wurde der kodierende Bereich der cDNA in 2 überlappenden Subfragmenten mittels PCR (XL62-6) mit Hilfe der Primer Nr. 240655 und Nr. 121339 für das N-terminale DNA-Fragment, sowie den Primern Nr. 240656 (enthält HindIII und Pmel Schnittstellen) und Nr. 121234 für das C-terminale DNA-Fragment amplifiziert. Die erhaltenen DNA-Fragmente (Subfragmente) wurden in den Vektor pCR21 kloniert. Die beiden

Subfragmente wurden komplett sequenziert und schließlich die vollständige H-SemaL cDNA durch Insertion eines 0.6kb C-terminalen SstI-HindIII Restriktions-Fragments in das mit den Restriktionsenzymen SstI und HindIII geschnittene, das N-terminale DNA-Fragment enthaltende Plasmid, hergestellt. Aus diesem Plasmid pCR2.1-H-SemaL (gemäß Sequenz in Tabelle 7, SEQ ID NO. 34) wurde das komplette Gen mittels der EcoRI-Schnittstelle (in pCR2.1) und HindIII-Schnittstelle (in Primer Nr. 240656, Tabelle 6) herausgeschnitten und in einen entsprechend geschnittenen, konstitutiven Expressionsvektor pCDNA3.1(-)MycHisA (Invitrogen) ligiert. Aus dem resultierenden rekombinanten Plasmid pCDNA3.1(-)H-SemaL-MycHisA (gemäß Sequenz in Tabelle 8) wurde das EcoRI-Apal Fragment (ohne Myc-His-Tag) herausgeschnitten und in den induzierbaren Vektor pIND ligiert (Ecdysone-Inducible Mammalian Expression System, Invitrogen), der zuvor ebenfalls mit EcoRI-Apal geschnitten worden war. Das rekombinante Plasmid wurde mit pIND-H-SemaL-EA (Sequenz gemäß Tabelle 11) bezeichnet. Ein EcoRI-Pmel-Fragment (mit Myc-His-Tag) aus pCDNA3.1(-)H-SemaL-Myc-HisA (Sequenz gemäß Tabelle 9) wurde in einen mit EcoRI-EcoRV geschnittenen Vektor pIND eingesetzt. Das rekombinante Plasmid wurde mit pIND-H-SemaL-EE (Sequenz gemäß Tabelle 10) bezeichnet. Ein Fusionsgen von H-SemaL mit Enhanced Green Fluorescent Protein (EGFP) wurde hergestellt durch Ligation des mit PCR amplifizierten EGFP-Leserahmens (aus dem Vektor pEGFP-C1 (Clontech), mit Hilfe der Primer Nr. 243068 + Nr. 243069, Taq52-60) in die Pmel-Schnittstelle des Plasmids pCDNA3.1(-)H-SemaL-MycHisA wodurch das Plasmid pCDNA3.1(-)H-SemaL-EGFP-MycHisA (Sequenz gemäß Tabelle 9) erhalten wurde.

In den Tabellen 7 bis 13 bedeuten kleine Buchstaben die Sequenz von H-SemaL, Teilen oder Derivaten derselben und große Buchstaben die Sequenz des Plasmids.

20 Beispiel 8:

Zur Herstellung von H-SemaL spezifischen Antikörpern wurden cDNA-Fragmente von H-SemaL in prokaryotische Expressionsvektoren integriert, in *E. coli* exprimiert und die Semaphorin-Derivate aufgereinigt. Die Semaphorin-Derivate wurden als Fusionsproteine mit einem His-Tag exprimiert. Dementsprechend wurden Vektoren verwendet, die die Sequenz für ein His-Tag enthalten und eine Integration des Semaphorin cDNA-Fragments im Leserahmen ermöglichen. Ein N-terminales 6xHistidin-Tag ermöglicht z.B. eine Aufreinigung mittels Nickel-Chelat-Affinitätschromatographie (Qiagen GmbH, Max-Volmer Straße 4, 40724 Hilden):

- 30 1. Der für die Aminosäuren 179-378 kodierende Teil der H-SemaL cDNA wurde mittels PCR mit den Primern Nr. 150788 und Nr. 150789 amplifiziert und dieses DNA-Fragment in den Vektor pQE30 (Qiagen), der zuvor mit den Restriktionsenzymen BamHI und HindIII geschnitten worden war, ligiert (Konstrukt pQE39-H-SemaL-BH (Sequenz gemäß Tabelle 12)).
- 35 2. Der für die C-terminalen Aminosäuren 480-666 kodierende Abschnitt der H-SemaL cDNA wurde mit den Restriktionsenzymen SstI und HindIII aus dem Plasmid pCR 2.1 geschnitten und in den Vektor pQE31 (Qiagen), der zuvor mit SstI und HindIII geschnitten worden war ligiert (Konstrukt pQE31-H-SemaL-SH (Sequenz gemäß Tabelle 13)).

40 Die korrekte Integration der Sequenzen im richtigen Leserahmen wurde durch DNA-Sequenzierung überprüft. Die Fusionsproteine, bestehend aus einem N-terminalen 6xHistidin-Taq und einem Teil des Semaphorins H-SemaL wurden mittels Ni²⁺-Affinitätschromatographie aufgereinigt. Die aufgereinigten Fusionsproteine wurden zur Immunisierung von verschiedenen Tieren (Hase, Huhn, Maus) benutzt.

45 Beispiel 9:

FACS-Analyse verschiedener Zelltypen (Figuren 4 und 5)
Die Zellen (ca. 0.2-0.5 x 10⁶) wurden mit FACS-Puffer gewaschen (Phosphate-buffered Saline (PBS) mit 5% fötalem Kälberserum (FCS) und 0.1% Na-Azid) und dann jeweils (auf Eis) für 1 Stunde mit den Antisera inkubiert. Als primäre Antikörper dienten für die Kontrolle (Overlay Hühner-Präimmunserum (1 : 50) und für den spezifischen Nachweis (Spezifische Färbung) ein H-SemaL spezifisches Hühner-Antiserum (1 : 50). Das spezifische Antiserum mit Antikörpern gegen Aminosäuren (Aa) 179-378 (mit N-terminalen His-Tag) von H-SemaL wurde durch Immunisierung von Hühnern mit dem durch Ni-Chelat-Affinitätschromatographie gereinigten Protein erzeugt (wie in Beispiel 8 beschrieben). Als zweiter Antikörper wurde ein FITC-markierter anti-Huhn F(ab') Antikörper aus Kaninchen verwendet (Dianova Jackson Laboratories, Best.-Nr. 303-095-006, Hamburg, Deutschland) (1mg/ml). Für die CD100-Färbung: wurde ein Rabbit-anti-Maus IgG, FITC-markiert verwendet. Der zweite Antikörper wurde jeweils in 1 : 50 Verdünnung in FACS-Puffer eingesetzt. Dann wurden die Zellen gewaschen, in PBS resuspendiert und im FACS analysiert. Die FACS-Analyse wurde mit

5 einem FACS-Trak Gerät (Becton—Dickinson) durchgeführt Prinzip: Eine Einzelzellsuspension wird in einen Messkanal vorbeigeleitet, dort werden die Zellen mit Laserlicht von 488nm bestrahlt und so Fluoreszenzfarbstoffe (FITC) angeregt. Gemessen werden Streulicht nach vorne (forward scatter, FSC: korreliert mit der Zellgröße), zur Seite (sideward scatter, SSC: korreliert mit dem Granulargehalt: bei unterschiedlichen Zelltypen unterschiedlich) und Fluoreszenz im Kanal 1 (FL 1) (für Wellenlängen im FITC Emissionsbereich, max. bei 530nm). Auf diese Weise wurden je 10000 Ereignisse (Zellen) gemessen.

10 Der Dotplot (Figuren 4a - k) (jeweils linke Abbildung): FSC gegen SSC (Größe gegen Granulargehalt/Streuung), darin eingegrenzt ist die im rechten Fenster (jeweils zugehörige rechte Abbildung) analysierte (einheitliche) Zellpopulation von ähnlicher Größe und Granulargehalt. Das rechte Fenster zeigt die Intensität von FL1 (X-Achse) gegen die Zahl der Ereignisse (Y-Achse), also eine Häufigkeitsverteilung.

15 Hierbei ist jeweils das Ergebnis mit dem Kontrollserum (nichtgefüllte Kurve) dem Ergebnis der spezifischen Färbung (ausgefüllte Kurve) überlagert. Eine Verschiebung der Kurve für die spezifische Färbung gegenüber der Kontrolle nach rechts entspricht einer Expression von H-SemaL in den entsprechenden Zellen. Je weiter die Verschiebung, desto stärker ist die Expression.

15 Für FACS Analyse verwendete Zelllinien:

- a) Zelllinie U937
American Type Culture Collection ATCC; ATCC Nummer: CRL-1593 Name: U-937
Gewebe: lymphoma; histiocytic; Monozytenähnlich
- 20 b) Zelllinie THP-1
ATCC Nummer: TIB-202
Gewebe: monocyte; acute monocytic leukemia
Species: human
Hinterleger: H. Koren
- c) Zelllinie K-562
ATCC Nummer: CCL-243
Gewebe: chronic myelogenous leukemia
Species: human
Hinterleger: S. Tsuchiya
- 25 d) Zelllinie L-428
DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSMZ Nr: ACC 197
Zelltyp: human Hodgkin's lymphoma
- e) Zelllinie Jurkat
DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSMZ Nr: ACC 282
Zelltyp: human T cell leukemia
- 30 f) Zelllinie Daudi
ATCC Number: CCL-213
Gewebe: Burkitt's lymphoma; B lymphoblast; B Zellen
Species : human
Hinterleger: G. Klein
- g) Zelllinie LCL
EBV-transformierte lymphoblastoide B-Zelllinie.
- 35 h) Zelllinie Jiyoye (P-2003)
ATCC Number: CCL-87
Gewebe: Burkitt's lymphoma; B Zellen; B lymphocyte
Species: human
Hinterleger: W. Henle
- i) CBL-Mix57
humane T-Zelllinie (isoliert aus Blut) transformiert mit rekombinantem H. Saimiri (Wild-type ohne Deletion)
- 40 j) CBL-Mix59
humane T-Zelllinie (isoliert aus Blut) transformiert mit H. Samiri (Deletion von ORF71).

55 55 Beispiel 10: Proteingel und Western-Blot

Sekretierbare humane SemaL (Aminosäuren 42-549 in Tabelle 4 (ohne Signalpeptid und ohne Transmembrandomäne) wurde in das Plasmid pMelBac-A (Invitrogen, De Sehulp, Leek, Niederlande, Cv 1950-20) kloniert und auf diese

Weise das Plasmid pMelBacA-H-SemaL (Länge 6622bp) erzeugt (Tabelle 15, Figur 8). Das H-SemaL Derivat wurde im Baculovirus-System (Bac-N-Blue, Invitrogen) exprimiert. Die Expression wurde in den Insekten-Eizellen abgeleiteten Zelllinien Sf9 (von *Spodoptera frugiperda*) und High Five™ (von *Trichoplusia ni*, U.S. Pat. No. 5,300,435, gekauft von Invitrogen) durch Infektion mit den rekombinanten, plaquegereinigten Baculoviren ausgeführt.

5 Die Expression wurde nach den Angaben des Herstellers durchgeführt.
Anschließend wurden die Proteine in einem Gel aufgetrennt und das H-SemaL Derivat im Western-Blot nachgewiesen. Die Detektion wurde mit dem H-SemaL spezifischen Hühnerantiserum (vergl. Beispiel 8 und Figur 7) (Verdünnung: 1:100) durchgeführt. Der spezifische Hühnerantikörper wurde mit anti-IgY-HRP Konjugat (Verdünnung: 1:3000, vom Kaninchen; Dianova Jackson Laboratories) nach Angaben des Herstellers nachgewiesen.

10 Beispiel 11: Herstellung von pMelBacA-HSEMAL

Der rekombinante Vektor (pMelBacA-HSEMAL, 6622bp) wurde hergestellt, indem ein entsprechendes DNA-Fragment, das für die Aminosäuren 42-649 von H-SemaL kodiert, in den Vektor pMelBacA (4,8 kb, Invitrogen) kloniert wurde (vergl. Annotation zu pMelBacA-H-SemaL). Die Klonierung erfolgte über BamHI und EcoRI in frame hinter die in dem Vektor vorliegende Signalsequenz („honeybee melittin signal sequence“). Ein entsprechendes H-SemaL DNA-Fragment wurde mit dem Primerpaar H-Sema 1 baculo 5' und H-Sema 1 baculo 3' amplifiziert.

15 Primer zur Amplifikation (TaKaRa Ex Taq-Polymerase) und Klonierung:
„H-Sema 1 baculo 5“, zur Amplifikation ohne Signalsequenz und zur Einführung einer BamHI-Schnittstelle
20 5'-CCGGATCCGCCAGGGCCACCTAACGGAGCGG-3' (SEQ ID NO: 43)
„H-Sema 1 baculo 3“, zur Amplifikation ohne Transmembrandomäne und zur Einführung einer EcoRI-Schnittstelle
5'-CTGAATTTCAGGAGGCCAGGGCACAGGCATG-3' (SEQ ID NO: 44).

25 Abbildungen:

Figur 1:
Gewebespezifische Expression von H-SemaL
30 A) Mehrfach-Gewebe Northern Blot (Clontech, Heidelberg, Deutschland). Von links nach rechts sind aufgetragen: Je 2 µg Poly-A-RNA pro Spur aus Milz, Thymus, Prostata, Hoden, Eierstöcken, Dünndarm, Dickdarm Mukosa, peripheren (Blut-) Leukozyten. Größenstandards sind markiert.

35 Die Blots wurden unter stringenten Bedingungen mit einer 800 Basenpaaren langen H-SemaL Probe hybridisiert.

Figur 2:
Schematische Darstellung der Klonierung der H-SemaL cDNA und der genomischen Organisation der H-SemaL kodierenden Sequenzen (H-SemaL Gen)
40 Oben: Lokalisation der EST-Sequenzen (Zugangsnummern; Lage der EST-Sequenzen ist relativ zur AHV-Sema Sequenz dargestellt).
Darunter: Amplifizierte PCR- und RACE-Produkte sowie die Position der cDNA Klonen im Bezug auf die Lokalisation in der vollständigen H-SemaL cDNA und dem offenen Leserahmen (ORF) für das kodierte Protein.
Unten: Relative Position der Exons im H-SemaL Gen im Bezug auf die genomische Sequenz.
45 Die Position der verwendeten Oligonukleotid Primer ist durch Pfeile angezeigt

Figur 3:

50 Phylogenetischer Baum: Erhalten durch mehrfaches Alignment der aufgeführten Semaphorin Sequenzen. Aufgrund der Gruppierung der Semaphorine in dem phylogenetischen Baum kann auf deren phylogenetische Verwandtschaft geschlossen werden.

Figur 4:
55 FACS Analyse der H-SemaL Expression in verschiedenen Zelllinien bzw. verschiedenen Zelltypen (vergl. Beispiel 8).

Figur 5:

Vergleichende Analyse der CD 100 und H-SemaL Expression (vergl. Beispiel 9).

5 Figur 6: Expression von sekretierbarem humanen SEMAL (H-SemaL) in HiFive und Sf9-Zellen (vergl. Beispiel 10).

(Aa 42-649 in pMelBac-A (Invitrogen) im Baculovirus-System (Bac-N-Blue, Invitrogen) Detektion mit spezifischem Hühner-Antiserum (1:100) und anti-IgY-HRP Konjugat (1:3000, vom Kaninchen, Jackson Lab.)

1,4,6 uninfizierte HiFive Zellen (serumfrei)

10 2,3,5,7,8 mit rekombinantem Baculovirus infizierte HiFive Zellen (serumfrei)

M Rainbow molecular weight marker (Amersham RPN756)

9,10 infizierte Sf9 Zellen (serumhaltiges Medium).

Figur 7: Spezifität des Antiseraums

15 Spuren 1-3: Huhn 1; Spuren 4-6: Huhn 2

Spuren 1 und 4: Präimmunserum

Spuren 2 und 5: 60. Immunisierungstag

Spuren 4 und 6: 105. Immunisierungstag

20 Immunisiert wurde mit den Aminosäuren 179-378 von H-SemaL (mit aminoterminalem His-Tag) (vergl. Beispiel 8, Punkt 1.)

Figur 8: Abbildung der Plasmidkarte von pMelBacA-HSEMAL.

25 Das rekombinante Plasmid wurde wie in Beispiel 11 beschrieben, hergestellt.

30

35

40

45

50

55

5
10
15
20
25
30
35
40
45
50

Tabelle 1: Verschiedene Subtypen von Semaphorinen aus verschiedenen Spezies

Name	Synonym	Spezies	Referenz
H-Sema III	(H-SemaD)	Mensch	Sez. (Kolodkin et al. 1993)
CD-100		Mensch	TM, IC; CD45 assoziiert, in T-Zellen exprimiert (Hall et al. 1996)
H-Sema V	(H-SemaA)	Mensch	Sez.; Locus 3p21.3 (Sekido et al. 1996; Roche et al. 1996)
H-Sema IV	(H-Sema 3F)	Mensch	Sez.; Locus 3p21.3 (Xiang et al. 1996; Sekido et al. 1996)
H-SemaE		Mensch	Sez.; am 3'Ende von M-SemaE divergierend (Leserahmen im Alignment verbessert) AB000220 (Yamada 1997 unpublished)
H-SemaK	KIAA0331	Mensch	Sez.; (Nagase et al. 1997)
H-SemaL	SEMA1	Mensch	TM, kein IC Diese Anmeldung (Püschel et al. 1995)
M-SemaA		Maus	Sez. (Püschel et al. 1995)
M-SemaB		Maus	TM, IC (Püschel et al. 1995)
M-SemaC		Maus	TM, IC (Püschel et al. 1995)
M-SemaD	M-Sema III	Maus	Sez. (MessenSmith et al. 1996; Püschel et al. 1995)
M-SemaE		Maus	Sez.; 5'partielle Sequenz (Püschel et al. 1995)
M-SemaF1	M-SemaF	Maus	TM, IC (Inagaki et al. 1995)
M-SemaG2	M-SemaG	Maus	TM, IC; exprimiert in Lymphoiden Zellen, Maus-Homolog zu CD100 (Funuyama et al. 1996)

5
10
15
20
25
30
35
40
45
50
55

Name	Synonym	Spezies		Referenz
M-Sema-F2	M-Sema-F	Maus	TM, IC; Thrombospondin-Motiv	(Adams et al. 1996)
M-SemaG1	M-SemaG	Maus	TM, IC; Thrombospondin-Motiv	(Adams et al. 1996)
M-SemaH		Maus	Sez.	(Christensen 1996 unpubl.) Z80941
M-SemaV/ia		Maus	TM, IC	(Zhou et al. 1997)
M-Semal	Semal	Maus	Partielle Sequenz	Diese Anmeldung
Collapsin-1		Huhn	Sez.	(Luo et al. 1993)
Collapsin-2		Huhn	Sez.	(Luo et al. 1995)
Collapsin-3		Huhn	Sez.	(Luo et al. 1995)
Collapsin-4		Huhn	Partielle Sequenz	(Luo et al. 1995)
Collapsin-5		Huhn	Sez.	(Luo et al. 1995)
R-Sema III		Ratte	Sez.	(Giger et al. 1996)
T-Sema I		Tribolium confusum	TM, IC	(Kolodkin et al. 1993)
Ce-Sema I		C.elegans	TM, IC	U15667 (Roy 1994 unpublished)
G-Sema I	Fascidin-IV	Grashüpfer	TM, IC	(Kolodkin et al. 1992)
D-Sema I		Drosophila	TM, IC	(Kolodkin et al. 1993)
D-Sema II		Drosophila	Sez.	(Kolodkin et al. 1993)
AHV-Sema		AHV-1	Sez.	(Ensser and Fleckenstein, 1995)

5

10

15

20

25

30

35

40

45

50

55

Name	Synonym	Spezies	Referenz
ORF-A39		Vaccinia Sez.	(Kolodkin et al. 1993)
ORF-A39-homolog		Variola Sez.;	(Kolodkin et al. 1993)

TM: Transmembrandomäne

Sez: sezemiert

IC: vermutlich intrazelluläres zyttoplasmatisches Sequenz-Motiv

Tabelle 2: cDNA-Sequenz von H-SemaL (2636 Nukleotide) (SEQ ID NO.: 1)

5	1	cggggccacg ggatgacgcc tcctccgccc ggacgtgccg ccccccagcgc
	51	accgcgcgcc cgcgccctg gcccgcggc tcggttgggg cttccgcgtgc
10	101	ggctgcggct gctgctgctg ctctggccgg ccggccgcctc cgcccaggc
	151	cacctaagga gcggaccccg catcttcgcc gcctggaaag gccatgttagg
15	201	gcaggaccgg gtggactttg gccagactga gcccacacg gtgcctttcc
	251	acgagccagg cagtcctct gtgtgggtgg gaggacgtgg caaggctcac
	301	ctctttact tccccgaggg caagaacca tctgtgcga cggtaatat
20	351	cggctccaca aaggggctct gtctggataa gcgggactgc gagaactaca
	401	tcacttcctt ggagaggcgg agtgaggggc tgctggcctg tggcaccaac
	451	gcccggcacc ccagctgctg gaacctggta aatggcactg tggtgccact
25	501	tggcgagatg agaggctacg cccccccttca gccggacgag aactccctgg
	551	ttctgtttga aggggacgag gtgtattcca ccattccggaa gcaggaatac
	601	aatgggaaga tccctcggtt ccgcccgcatt cggggcgaga gtgagctgta
	651	caccagtat actgtcatgc agaaccacca gttcatcaaa gccaccatcg
30	701	tgcaccaaga ccaggcttac gatgacaaga tctactactt cttccgagag
	751	gacaatcctg acaagaatcc tgaggctctt ctcaatgtgt cccgtgtggc
	801	ccagttgtgc aggggggacc agggtgggaa aagttcactg tcagtctcca
	851	agtggAACAC tttctgaaa gccatgctgg tatgcgtga tgctgcccacc
35	901	aacaagaact tcaacaggct gcaagacgtc ttcctgctcc ctgaccccaag
	951	cggccagtgg agggacacca gggctatgg tggtttctcc aacccttggaa
	1001	actactcagc cgctgtgtg tattccctcg gtgacattga caaggcttc
	1051	cgtacccctt cactcaaggg ctaccactca agccttccca acccgccggcc
40	1101	tggcaagtgc ctcccagacc agcagccat acccacagag accttccagg
	1151	tggctgaccg tcacccagag gtggcgacaa gggtggagcc catggggcct
	1201	ctgaagacgc cattgttcca ctctaaatac cactaccaga aagtggccgt
	1251	tcaccgcattt caagccagcc acggggagac ctttcatgtg ctttacctaa
	1301	ctacagacag gggcactatac cacaagggtgg tggaaccggg ggagcaggag
	1351	cacagcttcg cttcaacat catggagatc cagcccttcc gccgcgcggc
45	1401	tgcccatccag accatgtcgc tggatgtga gcgaggaaag ctgtatgtga
	1451	gctcccaagtgg ggggtgagc caggtgcggcc tggacactgtg tgaggctat
	1501	ggcgccggctt gccacgggtt cctcatgtcc cgagacccct actgcggctg
	1551	ggaccaggcc cgctgcattt ccatctacat ctccgaacgg tcagtgtgc
50	1601	aatccattaa tccagccgag ccacacaagg agtgccttccaa ccccaaacc
	1651	gacaaggcccc cactgcagaa gtttccctg gccccaaact ctcgtacta
	1701	cctgagctgc cccatgaaat cccggccacgc cacctactca tggcgccaca

1751 aggagaacgt ggagcagagc tgcgaacctg gtcaccagag ccccaactgc
 1801 atccctgttca tcgagaacct cacggcgcag cagtacggcc actacttctg
 1851 cgaggcccag gagggctcct acttccgcga ggctcagcac tggcagctgc
 1901 tgcccgagga cggcatcatg gcccagcacc tgctgggtca tgcctgtgcc
 1951 ctggctgcct ccctctggct gggggctg cccacactca ctcttggctt
 2001 gctggtccac tagggcctcc cgaggctggg catgcctcag gcttctgcag
 2051 cccagggcac tagaacgtct cacactcaga gcccggctggc ccgggagctc
 2101 cttgcctgcc acttcttcca ggggacagaaa taacccagtg gaggatgcca
 2151 ggcctggaga cgtccagccg caggcggctg ctggggccca ggtggcgcac
 2201 gnatggtgag gggctgagaa tgagggcacc gactgtgaag ctggggcatc
 2251 gatgacccaa gactttatct tctggaaaat atttttcaga ctccctcaaac
 2301 ttgactaaat gcagcgtatgc tcccagccca agagcccatg ggtcgaaaaag
 2351 tgggtttgga taggagagct gggactccat ctcgaccctg gggctgaggg
 2401 ctgagtcctt ctggactctt ggtacccaca ttgcctccctt cccctccctc
 2451 tctcatggct ggggtggctgg tggtcctgaa gaccggggc taccctctgt
 2501 ccagccctgt cctctgcagc tccctctctg gtcctgggtc ccacaggaca
 2551 gccgccttgc atgtttatttg aaggatgtt gctttccgga cgaaaggacg
 2601 gaaaaagctc tgaaaaaaaaaaaaaaaaaaaaaa
 26

Tabelle 3: Nukleotidsequenz der cDNA von M-Semal (partiell, 1195 Nukleotide)

30 (SEQ ID NO.: 2)

1 cggggctgctgatgacgccc tccctctccc ggacgtggccg ccccaagcgc
 51 accgcgcgcgc cgcgtctca gcctggccggc tcgggttcggg ctcccgctgc
 101 ggctcgccgtctcgtgggtt ttctgggtgg ccggccgcctc cggccaaaggc
 151 cactcgagga gcggaccccg catctccgcgt gtctggaaag ggcaggacca
 201 tgtggactttt agccagcctg agccacacac cgtgttttc catgagccgg
 251 gcagcttc tgcgtgggtt ggtggacgtg gcaaggctta ccacttcaac
 301 ttcccccagg gcaagaatgc ctctgtgcgc acggtaaca tcggctccac
 351 aaagggggtcc tgcaggaca aacaggactg tgggaattac atcactttc
 401 tagaaaggcg gggtaatggg ctgctgggtct gtggcaccaa tgccggaaag
 451 cccagctgtt ggaacttggt gaatgacagt gtgggtatgt cacttgggtga
 501 gatgaaaggc tatgccccct tcagccggat tgagaactcc ctgggtctgt
 551 ttgaaggaga tgaagtgtac tctaccatcc ggaaggcaggaa atacaacggg
 601 aagatccctc ggtttcgacg cattcggggc gagagtgaac tgtacacaag
 651 tgatacagtc atgcagaacc cacagttcat caaggccacc attgtgcacc
 701 aagaccaagc ctatgatgat aagatctact acttcttccg agaagacaac
 751 cctgacaaga accccggaggc tcctctcaat gtgtcccgag tagcccaagtt
 55

801 gtgcaggggg gaccaggggtg gtgagagtgc gttgtctgtc tccaaagtgg
 851 acaccccttccct gaaaggccatg ttggctctgca gcgatgcagc caccacacagg
 901 aacttcaatc ggctgcaaga tgtcttcctg ctccctgacc ccagtggcca
 951 gtggagagat accaggggtct atggcggttt ctccaaacccc tggaaactact
 1001 cagctgtctg cgtgtattcg cttgggtaca ttgacagagt cttccgtacc
 1051 tcatacgctca aaggctacca catggccctt tccaaaccctc gacctggcat
 1101 gtgcctccca aaaaagcagc ccataccac agaaaccttc caggtagctg
 1151 atagtcaccc agaggtggct cagaggggtgg aacctatggg gcccc

15 Tabelle 4: Aminosäuresequenz von H-SemaL (666 Aminosäuren)
(SEQ ID NO.: 3)

20 1 MTPPPPGRAA PSAPRARVPG PPARGLPLR LRLLLLLWAA AASAQGHLRS
 51 GPRIFAVWKG HVGQDRVDFG QTEPHVLFH EPGSSSVWVG GRGVYLFDF
 101 PEGKNASVRT VNIGSTKGSC LDKRDCENYI TLLERRSEGL LACGTNARHP
 151 SCWNLVNGTV VPLGEMRGYA PFPSPDENSLV LFEGDEVYST IRKQEYNGKI
 201 PRFRRIRGES ELYTSDTVMQ NPQFIKATIV HQDQAYDDKI YYFFREDNPD
 25 251 KNPEAPLNVS RVAQLCRGDQ GGESSLVSVK WNTFLKAMLV CSDAATNKNF
 301 NRLQDVFLLP DPSGQWRDTR VYGVFSNPWN YSAVCVYSLG DIDKVFRTSS
 351 LKGYHSSLPN PRPGKCLPDQ QPIPTETFQV ADRHPEVAQR VEPGPLKTP
 30 401 LFHSKYHYQK VAVHMRMQASH GETFHVLYL TDRGTTIHKVV EPGEQEHSFA
 451 FNIMEIQPFR RAAAIQTMISL DAERRKLYVS SQWEVSQVPL DLCEVYGGC
 501 HGCLMSRDPY CGWDQGRCIS IYSSERSVLQ SINPAEPHKE CPNPKPDKAP
 551 LQKVSLAPNS RYYLSCPME RHATYSWRHK ENVEQSCEPG HQSPNCILFI
 35 601 ENLTAQQYGH YFCEAQEGSY FREAQHWQLL PEDGIMAEHL LGHACALAAAS
 651 LWLGVLPTLT LGLLVH

40 Tabelle 5: (Partielle) Aminosäuresequenz von M-SemaL (394 Aminosäuren, entspricht Position 1-396 von H-SemaL) (SEQ ID NO.: 4)

45 1 MTPPPPGRAA PSAPRARVLS LPARFGLPLR LRLLLFWVA AASAQGHSRS
 51 GPRISAVWKG QDHVDFSQPE PHTVLFHEPG SFSVWVGGRG KVYHFNFPEG
 101 KNASVRTVNI GSTKGSCQDK QDCGNYITLL ERRGNGLLVC GTNARKPSCW
 151 NLVNDSVVMS LGEMKGYAPF SPDENSILVF EGDEVYSTIR KQEYNGKIPR
 50 201 FRRIRGESEL YTSDTVMQNP QFIKATIVHQ DQAYDDKIYY FFREDNPDKN
 251 PEAPLNVSRV AQLCRGDQGG ESSLVSVKW TFLKAMLVCS DAATNRNFNR

301 LQDVFLLPDP SGQWRDTRVY GVFSNPWNYS AVCVYSLGDI DRVFRSSLK
351 GYHMGLSNPR PGMCLPKKQP IPTETFQVAD SHPEVAQRVE PMGP

5

Tabelle 6: Synthetische Oligonukleotide, (Eurogentec, Seraing, Belgien)

10 Nummer des Primers/Bezeichnung	15 Nukleotidsequenz des Primers (der synthetischen Oligonukleotide)	20
91506/AP2	actcaactataggcgtcgagcggc	(SEQ ID NO.: 5)
121234	agccgcacacggtgctttc	(SEQ ID NO.: 6)
121235/Est 2	gcacagatgcgttcttgc	(SEQ ID NO.: 7)
121236/Est 3	accatagaccctggtgtccc	(SEQ ID NO.: 8)
121237/Est 4	gcagtgtatgctgccaccaac	(SEQ ID NO.: 9)
121238	ccagaccatgtcgctggatg	(SEQ ID NO.: 10)

25

30

35

40

45

50

55

121239/Est 6	acatgaggcaaccgtggcag	(SEQ ID NO.: 11)
131989/AP1	ccatcctaatacgaactcaactataggc	(SEQ ID NO.: 12)
5 131990/Est 7	aggtagaccccttgcacgtcc	(SEQ ID NO.: 13)
131991	gaacctcaacaggctgcaagacg	(SEQ ID NO.: 14)
131992	atgctgagcggaggaagctg	(SEQ ID NO.: 15)
131993	ccgcccatacacccacacacag	(SEQ ID NO.: 16)
10 150788	ctggaaagcttctgtggatcggtgc	(SEQ ID NO.: 17)
150789	tttggatccctggttctgtttgaag	(SEQ ID NO.: 18)
167579/cDNA		
ttctagaattcagcggccgtttttttttttttttttttttttvn		
15 Synthese Primer	(SEQ ID NO.: 19)	
168421	ggggaaagttcactgtcagtctccaag	(SEQ ID NO.: 20)
168422	ggaatacacacacagacggctgatag	(SEQ ID NO.: 21)
20 207608/	agcaagttcagcctggtaagt	(SEQ ID NO.: 22)
20 Amplifikation von λgt10 Insert		
207609/	ttatgagtatttcttccaggg	(SEQ ID NO.: 23)
Amplifikation von λgt 10 Insert		
25 232643/Est 13	ccattaatccagccgagccacacaag	(SEQ ID NO.: 24)
232644/Est 14	catctacagctccgaacggtcagtg	(SEQ ID NO.: 25)
233084	cagcggaaagcccaaccgag	(SEQ ID NO.: 26)
240655/hs 5	gggatgacgcctccgcggccgg	(SEQ ID NO.: 27)
30 240656/hs 3	aagcttacgtggaccagcaagccaagatg	(SEQ ID NO.: 28)
240657/hs 3c	aagcttttccgtccttccgtccgg	(SEQ ID NO.: 29)
243068	atggtagaccaaggcgcaggagctg	(SEQ ID NO.: 30)
243069	cttgcacagctcgccatgcccag	(SEQ ID NO.: 31)
35 260812	GGGTGGTGAGAGTCGTTGTCTGTC	(SEQ ID NO.: 32)
260813	GAGCGATGAGGTACCGAAGACTCTG	(SEQ ID NO.: 33)

Tabelle 7: Nukleotidsequenz des rekombinanten Plasmids pCR2.1-H-SemaL (SEQ ID NO.: 34)

40 1	AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA	
51	TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGGGCAA	
101	CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCAG GCTTTACACT	
45 151	TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATT	
201	CACACAGGAA ACAGCTATGA CCATGATTAC GCCaagcttc acgtggacca	
251	gcaagccaag agtgagtgtg ggcagcaccc ccagccagag ggaggcagcc	
301	aggcacagg catgaccagg caggtctcg gccatgatgc cgtcctcggg	
351	cagcagctgc cagtctgatg cctcgccgaa ttaggagccc tcctggcct	
401	cgcagaagta tgccgtac tgctgcgcgc tgaggtctc gatgaacagg	

451 atgcagttgg ggctctggcg accaggttcg cagctctgct ccacgttctc
 501 ctgtggcg catgagtagg tggcgtggcg ggattccatg gggcagctca
 551 gtagtagcg agagttggg gccaggaaa cttctgcag tggggccttg
 601 tctggttgg gggtggaca ctccctgtgt ggctcggctg gattaatgga
 651 ttgcagcact gaccgttcgg agctgttagat ggagatgcag cggccctgg
 701 cccagccgca gttagggtct cgggacatga ggcaaccgtg gcagccccgg
 751 ccatagaccc cacacaggc caggggcacc tggctcacct cccactgg
 801 gctcacatac agcttcctcc gctcagcatc cagcgacatg gtctggatgg
 851 cagccgcgcg gcggaaaggc tggatctcca tggatgtgaa ggcgaagctg
 901 tgctcctgct ccccggttc caccaccttgg tggatagtgc ccctgtctgt
 951 agtaggtaa agcacatgaa aggtctcccc gtggctggct tgcacatgcgg
 1001 gaacggccac ttctggtag tggatttag agtggaaacaa tggcgtctc
 1051 agaggccccca tgggctccac cctctgcgc acctctgggt gacggtcagc
 1101 caccttggaaag gtctctgtgg gtatcggctg ctggtctggg aggaacttgc
 1151 caggccgcgg gttggaaagg cttgagtggt agcccttgag tgaggaggta
 1201 cgaagacact tgcataatgtc accggggaa tacacacaga cggctgagta
 1251 gtccagggg ttggagaaaa caccatagac cctggtgcc ctccactggc
 1301 cgctggggtc agggagcagg aagacgtctt gcagcccttggt gaagttctt
 1351 ttggtggcag catcaactgca taccacatg gcttcagaa aagtgttcca
 1401 ctggagact gacagtgaac ttccccacc ctggtcccccttgcacaact
 1451 gggccacacg ggacacatttgg agaggagcct caggatttttgcaggattt
 1501 tcctctcgga agaagtagta gatctgtca tgcgttggctt ggtcttgg
 1551 cacgatggtg gctttgatga actgtgggtt ctgcataatgca gtatcactgg
 1601 tgtagacttgc actctcgccc cggatgcggc ggaaccgagg gatcttcca
 1651 ttgtattccct gcttccggat ggttggataac acctcgcccccttcaaaacag
 1701 aaccaggag ttctcgccg ggctgaagg ggcgttggctt ctcatctcg
 1751 caagtggcac cacagtgcac ttaccatgg tccacatgg ggggtggccgg
 1801 gcgttggtgc cacaggccag cagccccctca ctccgcctt ccaggagagt
 1851 gatgttagttc tcgcagttccc gcttatccag acaggacccc tttgtggagc
 1901 cgatattcac cgatgcacaca gatgcgttct tggccctggg gaagtcaaaag
 1951 agtagacact tgcacatgtcc tccacccac acagaggagc tgcctggctc
 2001 gtggaaaagc accgtgtgcg gctcagttcg gccaaagtcc acccggttcc
 2051 gcccataatg gccttccag acggcgaaga tgcggggtcc gtcctttagg
 2101 tggccctggg cggaggccgcg ggcccccag agcagcagca gcagccgcag
 2151 ccgcagcggaa agcccccaacc gagccggcg gccaggacg cggccgcgcg
 2201 gtgcgttggg ggcggcacgt ccggggcgag gaggcgtcat cccaaagccga
 2251 attcTGCAGA TATCCATCAC ACTGGCGGCC GCTCGAGCAT GCATCTAGAG
 2301 GGCCCAATTC GCCCTATAGT GAGTCGTATT ACAATTCACT GGCGTGTGTT
 2351 TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT

2401 TGCAGCACAT CCCCTTTGCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA
 2451 CCGATGCCCG TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACCGG
 5 2501 CCCTGTAGCG GCGCATTAAG CGCGCGGGGT GTGGTGGTTA CGCGCAGCGT
 2551 GACCGCTACA CTTGCCAGCG CCCTAGGCC CGCTCCTTTC GCTTCTTCC
 2601 CTTCTTCTCG CGCCACGTTG GCCGGCTTTC CCCGTCAGC TCTAAATCGG
 2651 GGGCTCCCTT TAGGGTCCG ATTTAGAGCT TTACGGCACC TCGACCGCAA
 10 2701 AAAACTGAT TTGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA
 2751 CGGTTTTTCG CCCTTGACG TTGGAGTCCA CGTTCTTAA TAGTGGACTC
 2801 TTGTTCCAAA CTGGAACAAC ACTCAACCT ATCGCGGTCT ATTCTTTGA
 2851 TTTATAAGGG ATTTTGCAGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA
 15 2901 TTTAACAAAT TCAGGGCGCA AGGGCTGCTA AAGGAACCGG AACACGTAGA
 2951 AAGCCAGTCC GCAGAAACGG TGCTGACCCC GGATGAATGT CAGCTACTGG
 3001 GCTATCTGGA CAAGGGAAAA CGCAAGCGCA AAGAGAAAGC AGGTAGCTTG
 3051 CAGTGGGCTT ACATGGCGAT AGCTAGACTG GGCGGTTTA TGGACAGCAA
 20 3101 GCGAACCGGA ATTGCCAGCT GGGGCGCCCT CTGGTAAGGT TGGGAAGCCC
 3151 TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGGATCT GATGGCGCAG
 3201 GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG
 3251 AACAAAGATGG ATTGCACGCA GGTTCTCCGG CCGCTGGGT GGAGAGGCTA
 25 3301 TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT
 3351 GTTCCGGCTG TCAGCGCAGG GGCGCCCGGT TCTTTTGTC AAGACCGACC
 3401 TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG
 3451 CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA
 3501 AGCGGGAAAGG GACTGGCTGC TATTGGCGA AGTGCCGGGG CAGGATCTCC
 3551 TGTCATCTCG CCTTGCTCT GCCGAGAAAG TATCCATCAT GGCTGATGCA
 3601 ATGCGGCCGC TGCAACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA
 3651 AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG
 3701 TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCGAA
 3751 CTGTCGCCA GGCTCAAGGC GCGCATGCC GACGGCGAGG ATCTCGTCGT
 3801 GATCCATGGC GATGCCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT
 3851 TTTCTGGATT CAACGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG
 3901 GACATAGCGT TGGATACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG
 3951 GGCTGACCGC TTCCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTGCGAGC
 4001 GCATGCCCTT CTATGCCCTT CTTGACGAGT TCTTCTGAAT TGAAAAAGGA
 4051 AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTGCG
 4101 GCATTTGCC TTCCGTGTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA
 4151 AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC
 4201 TCAACAGCGG TAAGATCCTT GAGAGTTTC GCCCCGAAGA ACGTTTCCA
 4251 ATGATGAGCA CTTTAAAGT TCTGCTATGT CATAACTAT TATCCCGTAT
 4301 TGACGCCGGG CAAGAGCAAC TCGGTGCCG GGCGCGGTAT TCTCAGAATG

4351 ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG
 4401 ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC
 5 4451 GGCCAACCTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT
 4501 TTTTGCACAA CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG
 4551 GAGCTGAATG AAGCCATACC AAACGACGAG AGTGACACCA CGATGCCTGT
 4601 AGCAATGCCA ACAACGTTGC GCAAACATTAACTGGCGAA CTACTTACTC
 10 4651 TAGCTTCCCG GCAACAATTAA TAGACTGGTA TGAGGGCGGA TAAAGTTGCA
 4701 GGACCACTTC TGCGCTCGGC CCTTCCGGCT GGCTGGTTA TTGCTGATAA
 4751 ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC
 4801 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG
 15 4851 GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT
 4901 GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA
 4951 TTGATTTAAA ACTTCATTAA TAATTTAAAA GGATCTAGGT GAAGATCCTT
 20 5001 TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTT CGTCCACTG
 5051 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTT
 5101 TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG
 5151 GTGGTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTC CGAAGGTAAC
 25 5201 TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCCTTCTA GTGTAGCCGT
 5251 AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC ATACCTCGCT
 5301 CTGCTTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
 5351 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG
 30 5401 GCTGAACGGG GGGTTCGTGC ACACAGCCCA CCTTGGAGCG AACGACCTAC
 5451 ACCGAACCTGA GATACTTACA GCGTGAGCAT TGAGAAAGCG CCACGCTTCC
 5501 CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG
 5551 GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACCGCTGGTA TCTTTATAGT
 35 5601 CCTGTGGGT TTGCCCCACT CTGACTTGAG CGTCGATTTT TGTGATGCTC
 5651 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC
 5701 GGTTCCCTGGC CTTTGCTGG CCTTTGCTC ACATGTTCTT TCCTGCGTTA
 5751 TCCCCTGATT CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC
 40 5801 CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
 5851 CGGAAG

45 Tabelle 8: Nukleotidsequenz des rekombinanten Expressionsplasmids pCDNA3.1(-)H-Semal-MychIsA
 (SEQ ID NO.: 35)

50 1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG

55

151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTGCG
 201 CTGCTTCGCG ATGTACGGGC CAGATATACTG CGTTGACATT GATTATTGAC
 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTCAT AGCCCATATA
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
 351 CCCAACGACC CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
 401 AACGCCAATA GGGACTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
 451 AAAC TGCCCCA CTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA
 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
 601 TCGCTATTAC CATGGTGTATG CGGTTTGGC AGTACATCAA TGGGCGTGGA
 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA
 751 ACAACTCCGC CCCATTGACG CAAATGGCG GTAGGGTGT ACGGTGGGAG
 801 GTCTATATAA GCAGAGCTCT CTGGCTAATC AGAGAACCCCA CTGCTTACTG
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 901 GTTTAAACGG GCCCTCTAGA CTCGAGGGC CGCCACTGTG CTGGATATCT
 951 GCAaattcg gcttggatg acgcctccctc cggccggacg tgccggccccc
 1001 agcgacccgc ggcggccggt ccctggcccg cggcgtcggt tggggcttcc
 1051 gctgcccgtg cgctgctgc tgctgctctg ggcggccgccc gcctccggcc
 1101 agggccacct aaggagcggc ccccgcatct tcgcccgtctg gaaaggccat
 1151 gtagggcagg accgggtgga ctttggccag actgagccgc acacggtgct
 1201 tttccacgag ccaggcagct cctctgtgtg ggtggagga cgtggcaagg
 1251 tctacccctt tgacttcccc gagggcaaga acgcacatctgt ggcacggtg
 1301 aatatcggt ccacaaagggt gtcctgtctg gataagcggg actgcaagaa
 1351 ctacatcaact ctccctggaga ggcggagtg gggctgtctg gcctgtggca
 1401 ccaacgcccgc gcaccccccacg tgctggaaacc tggtaatgg cactgtggtg
 1451 ccacttggcg agatgagagg ctacgcccccc ttcaagccgg acgagaactc
 1501 cctgggtctg tttgaaggggg acgaggtgtt ttccaccatc cggaaagcagg
 1551 aatacaatgg gaagatccct cgggtccgccc gcattccgggg cgagagtgag
 1601 ctgtacacca gtgatactgt catgcagaac ccacagtca tcaaagccac
 1651 catcggtgcac caagaccagg cttacgatga caagatctac tacttcttcc
 1701 gagaggacaa tcctgacaag aatcctgagg ctcctctcaa tgtgtcccg
 1751 gtggcccaagt tgtgcagggg ggaccaggtt gggaaaagtt cactgtcagt
 1801 ctccaaagtgg aacacttttc tgaaagccat gctggtatgc agtgatgtcg
 1851 ccaccaacaa gaacttcaac aggctgcaag acgtttccct gctccctgac
 1901 cccagcggcc agtggagggc caccagggtc tatgggttt tctccaaccc
 1951 ctggaaactac tcagccgtct gtgtgttcc cctcggtgac attgacaagg
 2001 tcttccgtac ctccctcaactc aagggttacc actcaaggct tcccaaccc
 2051 cggcctggca agtgcctccc agaccagcag ccgataacccca cagagacctt

2101 ccaggtggct gaccgtcacc cagaggtggc gcagagggtg gagcccatgg
 2151 ggcctctgaa gacgccattg ttccactcta aataccacta ccagaaagtg
 5 2201 gccgttcaacc gcatgcaagc cagccacggg gagaccttc atgtgctta
 2251 cctaactaca gacagggca ctatccacaa ggtggggaa ccgggggagc
 2301 aggagcacag cttcgcccttc aacatcatgg agatccagcc cttccgccc
 2351 gcccgtgcca tccagaccat gtcgctggat gctgagcgg a ggaagctgta
 10 2401 tgtgagctcc cagtgggagg tgagccagg gcccctggac ctgtgtgagg
 2451 tctatggcg gggctgccc ac gttgcctca tgtccgaga cccctactgc
 2501 ggctgggacc agggccgctg catctccatc tacagctccg aacggtcagt
 2551 gctgcaatcc attaatccag ccgagccaca caaggagtgt cccaaacccca
 15 2601 aaccagacaa ggccccactg cagaagg ttt ccctggcccc aaactctcgc
 2651 tactacctga gtcgccccat ggaatccgc cacgcccc ac tccatggcg
 2701 ccacaaggag aacgtggagc agagctgca acctggtcac cagagcccc
 2751 actgcaccc tttcatcgag aacccacgg cgcagcagta cggccactac
 2801 ttctgcgagg cccaggaggg ctccacttc cgcgaggctc agcactggca
 2851 gctgctgccc gaggacggca tcatggccga gcacctgctg ggtcatgcct
 2901 gtgcctgga tgcctccctc tggctggggg tgctgcccac actcactctt
 2951 ggcttgcgg tccacgtgaa gcttGGGGCC GAACAAAAAC TCATCTCAGA
 25 3001 AGAGGATCTG AATAGCGCCG TCGACCATCA TCATCATCAT CATTGAGTTT
 3051 AAACCGCTGA TCAGCCTCGA CTGTGCCTTC TAGTTGCCAG CCATCTGTTG
 3101 TTTGCCCTC CCCCCGTGCCT TCCTTGACCC TGGAAAGGTGC CACTCCCAC
 3151 GTCCCTTCCT AATAAAATGA GGAAATTGCA TCGCATTGTC TGAGTAGGTG
 3201 TCATTCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG GGGGAGGATT
 3251 GGGAAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGCTC TATGGCTTCT
 3301 GAGGCGGAAA GAACCAGCTG GGCTCTAGG GGGTATCCCC ACGGCCCTG
 3351 TAGCGGCCGA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG
 3401 CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC
 3451 TTTCTCGCCA CGTTGCCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGCAT
 3501 CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC
 40 3551 TTGATTAGGG TGATGGTCA CGTAGTGGGC CATGCCCTG ATAGACGGTT
 3601 TTTGCCCTT TGACGTTGGA GTCCACGTTTC TTTAATAGTG GACTCTGTT
 3651 CCAAACACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTAT
 3701 AAGGGATTAA GGGGATTTG GCCTATTGGT TAAAAAAATGA GCTGATTAA
 45 3751 CAAAAATTAA ACGCGAATTA ATTCTGTGGA ATGTGTGTCA GTTAGGGTGT
 3801 GGAAAGTCCC CAGGCTCCCC AGGCAGGCAG AAGTATGCAA AGCATGCATC
 3851 TCAATTAGTC AGCAACCAGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC
 3901 AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC
 3951 CCTAACTCCG CCCATCCCCC CCCTAACTCC GCCCCAGTTCC GCCCATTCTC
 4001 CGCCCCATGG CTGACTAATT TTTTTATTT ATGCAGAGGC CGAGGCCGCC

4051 TCTGCCTCTG AGCTATTCCA GAAAGTAGTGA GGAGGCTTT TTGGAGGCCT
 4101 AGGCTTTGC AAAAGCTCC CGGGAGCTT GATATCCATT TTCGGATCTG
 4151 ATCAAGAGAC AGGATGAGGA TCGTTCGCA TGATTGAACA AGATGGATTG
 4201 CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTG GCTATGACTG
 4251 GGCACAAACAG ACAATCGGCT GCTCTGATGC CGCGTGTTC CGGCTGTCA
 4301 CGCAGGGCGC CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG
 4351 AATGAACCTGC AGGACGAGGC AGCGCGGCTA TCGTGGCTGG CCACGACGGG
 4401 CGTCCCTTGC GCAGCTGTGC TCGACGTTGT CACTGAAGCG GGAAGGGACT
 4451 GGCTGCTATT GGGCGAAGTG CGGGGGCAGG ATCTCCTGTC ATCTCACCTT
 4501 GCTCCTGCCG AGAAAAGTATC CATCATGGCT GATGCAATGC GGCAGCTGCA
 4551 TACGCTTGAT CCGGCTACCT GCCCCATTGCA CCACCAAGCG AAACATCGCA
 4601 TCGAGCGAGC ACCTACTCGG ATGGAAGCCG GTCTGTGCA TCAGGGATGAT
 4651 CTGGACGAAG AGCATCAGGG GCTCGCGCA GCCGAACCTGT TCGCCAGGCT
 4701 CAAGGCGCGC ATGCCCAGC GCGAGGATCT CGTCGTGACC CATGGCGATG
 4751 CCTGCTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTTCATC
 4801 GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC
 4851 TACCCGTGAT ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC
 4901 TCGTGCCTTA CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT
 4951 CGCCTTCTTG ACGAGTTCTT CTGAGCGGGGAA CTCTGGGGTT CGAAATGACC
 5001 GACCAAGCGA CGCCCAACCT GCCATCACGA GATTCGATT CCACCGCCGC
 5051 CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA
 5101 TGATCCTCCA GCGCGGGGAT CTCATGCTGG AGTTCTCGC CCACCCCAAC
 5151 TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA
 5201 TTTCACAAAT AAAGCATTTC TTTCACTGCA TTCTAGTTGT GGTTTGTCCA
 5251 AACTCATCAA TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG
 5301 AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC
 5351 GCTCACAATT CCACACAACA TACGAGCGGG AAGCATAAAG TGAAAGCCT
 5401 GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG
 5451 CCCGCTTCC AGTCGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG
 5501 CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGGCCTCT TCCGCTTCC
 5551 CGCTCACTGA CTCGCTGCGC TCGGTGCGTTC GGCTGCGGGCG AGCGGTATCA
 5601 GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCGAG GGGATAACCC
 5651 AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
 5701 AGGCCGCGTT GCTGGCGTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT
 5751 CACAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA
 5801 AAGATACCAAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCAGC TCTCCTGTTC
 5851 CGACCCCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC
 5901 GTGGCGCTTT CTCAATGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT
 5951 CGTTCGCTCC AAGCTGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACCC

6001 GCTGCGCCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC
 6051 GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG
 5 6101 GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT
 6151 ACACCTAGAAG GACAGTATTG GGTATCTGCG CTCTGCTGAA GCCAGTTACC
 6201 TTCGGAAAAA GAGTTGGTAG CTCTGATCC GGCAAACAAA CCACCGCTGG
 6251 TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG
 10 6301 GATCTCAAGA AGATCCTTG ATCTTTCTA CGGGGCTGAA CGCTCAGTGG
 6351 AACGAAAAGT CACGTTAAGG GATTTGGTC ATGAGATTAT CAAAAGGAT
 6401 CTTCACCTAG ATCCTTTAA ATAAAAATG AAGTTTAAA TCAATCTAAA
 6451 GTATATATGA GTAAACCTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG
 15 6501 GCACCTATCT CAGCGATCTG TCTATTCGT TCATCCATAG TTGCCTGACT
 6551 CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
 6601 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA
 6651 GCAATAAACCG AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCCTGCAAC
 20 6701 TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA
 6751 GTAGTTGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC
 6801 ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC
 6851 CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
 6901 TTAGCTCCTT CGGTCCCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG
 6951 TTATCACTCA TGGTTATGCC AGCACTGCAT AATTCTCTTA CTGTCATGCC
 7001 ATCCGTAAGA TGCTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT
 7051 GAGAATAGTG TATGCCCGA CCGAGTTGCT CTTGCCGGC GTCAATACGG
 7101 GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA
 7151 ACGTTCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
 7201 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTACT
 7251 TTCACCAAGCG TTTCTGGGTG AGCAAAACCA GGAAGGCAAA ATGCCGAAA
 7301 AAAGGGATAA AGGGCGACAC GGAAATGTTG AATACTCATA CTCTCCCTT
 7351 TTCAATATTA TTGAACCAT TATCAGGGTT ATTGTCTCAT GAGCGGATAC
 7401 ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTTC CGCGCACATT
 7451 TCCCCGAAAAA GTGCCACCTG ACGTC

45 Tabelle 9: Nukleotidsequenz des rekombinanten Plasmids pcDNA 3.1-H-SemaL-EGFP-MychisA
 (SEQ ID NO.: 36)

1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
 50 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
 151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTGCG
 201 CTGCTTCGCG ATGTACGGGC CAGATATAACG CGTTGACATT GATTATTGAC

5 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCCATATA
301 TGGAGTTCGG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
351 CCCAACGACC CCCGCCATT GACGTCATAA ATGACGTATG TTCCCATACT
401 AACCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTACGGT
451 AACTGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC
501 CCTATTGACG TCAATGACGG TAAATGGCC GCCTGGCATT ATGCCAGTA
551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTC
601 TCGCTATTAC CATGGTGATG CGGTTTGGC AGTACATCAA TGGCGTGG
651 TAGGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
701 TGGGAGTTTG TTTGGCACCC AAAATCAACG GGACTTTCCA AAATGTCGTA
751 ACAACTCCGC CCCATTGACG CAAATGGCG GTAGGCGTGT ACGGTGGGAG
801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
901 GTTTAACCGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT
951 GCAaaattcg gcttggatg acgcctctc cggccggacg tgccgcccc
1001 agcgcaccgc gcgcggcggt ccctggcccg cggcgtcggt tggggcttcc
1051 gctgcggctg cggcgtctgc tgctgtctg ggcggccgccc gcctccggcc
1101 agggccacct aaggagcggg ccccgcatct tcgcccgtctg gaaaggccat
1151 gtagggcagg accgggtggg ctttggccag actgagccgc acacgggtct
1201 tttccacgag ccaggcagct ccctgtgtg gttgggagga cgtggcaagg
1251 tctacctctt tgacttcccc gagggaaga acgcacatctgt gcgcacggtg
1301 aatatcggtt ccacaaaggg gtcctgtctg gataagcggg actgcgagaa
1351 ctacatcaact ctccctggaga ggcggagtga ggggctgtctg gcctgtggca
1401 ccaacgcccc gcaccccgac tgctggaaacc tggtaatgg cactgtggtg
1451 ccacttggcg agatgagagg ctacgcccccc ttcaagcccg acgagaactc
1501 cctgttctg tttgaagggg acgagggtta ttccaccatc cggaaagcagg
1551 aatacaatgg gaagatccct cgggtccgccc gcatccgggg cgagagtgg
1601 ctgtacacca gtgatactgt catgcagaac ccacagttca tcaaagccac
1651 catcgtgcac caagaccagg cttacgtga caagatctac tacttcttcc
40 1701 gagaggacaa tcctgacaag aatcctgagg ctccctctcaa tgggtcccg
1751 gtggcccgat tggcagggg ggaccagggt gggaaaagtt cactgtcgat
1801 ctccaagtgg aacacttttc tggaaagccat gctggatgtc agtgtatgt
1851 ccaccaacaa gaaacttcaac aggctgcaag acgtcttccct gctccctgac
1901 cccagcggcc agtggaggga caccagggtc tatgggtttt tctccaaccc
1951 ctggaaactac tcagccgtct gtgtgtattc cctccgggtgac attgacaagg
2001 tcttccgtac ctccctcaactc aagggttacc actcaagcct tcccaacccg
2051 cggccctggca agtgcctccc agaccaggcag ccgataccca cagagacctt
2101 ccagggtggct gaccgtcacc cagagggtggc gcagagggtg gagcccatgg
2151 ggcctctgaa gacgccattt ttccactcta aataccacta ccagaaaatgt

2201 gccgttccacc gcatgcaagc cagccacggg gagaccttgc atgtgcttta
 2251 ccttaactaca gacaggggca ctatccacaa ggtgggtggaa ccggggggagc
 2301 aggagcacag cttccgccttc aacatcatgg agatccagcc cttccgcgc
 2351 gcggctgcca tccagaccat gtcgctggat gctgagcggaa ggaagctgt
 2401 tgtgagctcc cagtgggagg tgagccaggt gcccctggac ctgtgtgagg
 2451 tctatggcg 999ctgccc ggttgcctca tgtcccgaga cccctactgc
 2501 ggctgggacc agggccgctg catctccatc tacagctccg aacggtcagt
 2551 gctgaatcc attaatccag ccgagccaca caaggagtgt cccaaacccca
 2601 aaccagacaa ggccccactg cagaagggtt ccctggggcc aaactctcgc
 2651 tactacctga gctgccccat ggaatcccgc cacgcccaccc actcatggcg
 2701 ccacaaggag aacgtggagc agagctgcga acctggtcac cagagcccca
 2751 actgcacccct gttcatcgag aacctcacgg cgccagcgtt cggccactac
 2801 ttctgcgagg cccaggaggg ctccctacttc cgccaggctc agcaactggca
 2851 gctgctgccc gaggacggca tcatggccga gcacctgctg ggtcatgcct
 2901 gtgcccggc tgcctccctc tgctgggggg tgctgccccacttactctt
 2951 ggcttgctgg tccacATGGT GAGCAAGGGC GAGGAGCTGT TCACCGGGGT
 3001 GGTGCCCATC CTGGTCGAGC TGGACGGCGA CGTAAACGGC CACAAGTTCA
 3051 GCGTGTCCGG CGAGGGCGAG GGCAGATGCCA CCTACGGCAA GCTGACCCCTG
 3101 AAGTTCATCT GCACCACCGG CAAGCTGCCA GTGCCCTGGC CCACCCCTCGT
 3151 GACCACCCCTG ACCTACGGCG TGCACTGCTT CAGCCGCTAC CCCGACCACA
 3201 TGAAGCAGCA CGACTTCTTC AAGTCCGCCA TGCCCGAAGG CTACGTCCAG
 3251 GAGCGCACCA TCTTCTTCAA GGACGACGGC AACTACAAGA CCCGCGCCGA
 3301 GGTGAAGTTC GAGGGCGACA CCCTGGTGAA CCGCATCGAG CTGAAGGGCA
 3351 TCGACTTCAA GGAGGGACGGC AACATCCTGG GGCACAAAGCT GGAGTACAAC
 3401 TACAACAGCC ACAACGTCTA TATCATGGCC GACAAGCAGA AGAACGGCAT
 3451 CAAGGTGAAC TTCAAGATCC GCCACAACAT CGAGGACGGC AGCGTGCAGC
 3501 TCGCCGACCA CTACCAGCAG AACACCCCCA TCGGCAGCGG CCCCCGTGCTG
 3551 CTGCCCCGACA ACCACTACCT GAGCACCCAG TCCGCCCCCTGA GCAAAGACCC
 3601 CAACGAGAAG CGCGATCACA TGGTCCTGCT GGAGTTCGTG ACCGCCGCCG
 3651 GGATCACTCT CGGCATGGAC GAGCTGTACA Aggtgaagct tGGGGCCCGAA
 3701 CAAAAAACTCA TCTCAGAAGA GGATCTGAAT AGCGCCGTG ACCATCATCA
 3751 TCATCATCAT TGAGTTTAAAC CCGCTGATCA GCCTCGACTG TGCCTTCTAG
 3801 TTGCCAGCCA TCTGTTGTTT GCCCCCTCCCC CGTGCCTTCC TTGACCCCTGG
 3851 AAGGTGCCAC TCCCACGTGTC CTTTCCATAAT AAAATGAGGA AATTGCATCG
 3901 CATTGTCTGA GTAGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA
 3951 CAGCAAGGGG GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCGG
 4001 TGGGCTCTAT GGCTTCTGAG GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG
 4051 TATCCCCACG CGCCCTGTAG CGGGCGATTA AGCGCGGCCGG GTGTGGTGGT
 4101 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT

4151 TCGCTTCCTT CCCTTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA
 4201 GCTCTAAATC GGGGCATCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA
 5 4251 CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT
 4301 CGCCCTGATA GACGGTTTT CGCCCTTGA CGTTGGAGTC CACGTTCTT
 4351 AATAGTGGAC TCTTGTCCA AACTGGAACA AACTCAACC CTATCTCGGT
 4401 CTATTCTTT GATTATAAG GGATTTGGG GATTTCGGCC TATTGGTTAA
 10 4451 AAAATGAGCT GATTAAACAA AAATTTAACG CGAATTAATT CTGTGGAATG
 4501 TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG
 4551 TATGCAAAGC ATGCATCTCA ATTAGTCAGC ACCAGGTGT GGAAAGTCCC
 4601 CAGGCTCCCC AGCAGGCAGA AGTATGCAA GCATGCATCT CAATTAGTCA
 15 4651 GCAACCATAG TCCCAGCCCC AACTCCGCCC ATCCCAGCCCC TAACTCCGCC
 4701 CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG
 4751 CAGAGGCCGA GGGCGCCTCT GCCTCTGAGC TATTCCAGAA GTAGTGAGGA
 4801 GGCTTTTTG GAGGCCTAGG CTTTGCAA AAGCTCCGG GAGCTTGAT
 20 4851 ATCCATTTTC GGATCTGATC AAGAGACAGG ATGAGGATCG TTTCGCATGA
 4901 TTGAACAAGA TGGATTGCAC GCAGGTTCTC CGGGCGCTTG GGTGGAGAGG
 4951 CTATTCGGCT ATGACTGGGC ACAACAGACA ATCGGCTGCT CTGATGCCGC
 5001 CGTGTCCGG CTGTCAGCGC AGGGGCGCCC GGTCTTTTT GTCAAGACCG
 25 5051 ACCTGTCCGG TGCCCTGAAT GAACTGCAGG ACGAGGCAGC CGGGCTATCG
 5101 TGGCTGGCCA CGACGGCGT TCCTTGCAGCA GCTGTGCTCG ACGTTGTAC
 5151 TGAAGCGGGA AGGGACTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC
 30 5201 TCCTGTCAATC TCACCTTGCT CCTGCCGAGA AAGTATCCAT CATGGCTGAT
 5251 GCAATGCGGC GGCATGCATAC GCTTGATCCG GCTACCTGCC CATTGACCA
 5301 CCAAGCGAAA CATCGCATCG AGCGAGCACG TACTCGGATG GAAGCCGGTC
 5351 TTGTGATCA GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC
 35 5401 GAACTGTTCG CCAGGCTCAA GGCGCGCATG CCCGACGGCG AGGATCTCGT
 5451 CGTGACCCAT GGCATGCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC
 5501 GCTTTCTGG ATTCAATCGAC TGTGGCCGGC TGGGTGTGGC GGACCGCTAT
 5551 CAGGACATAG CGTTGGCTAC CGTGATATT GCTGAAGAGC TTGGCGGCGA
 40 5601 ATGGGCTGAC CGCTTCTCG TGCTTTACGG TATCGCCGCT CCCGATTCCG
 5651 AGCGCATCGC CTTCTATCGC CTTCTGACG AGTTCTCTG AGCGGGACTC
 5701 TGGGGTTCGA AATGACCGAC CAAGCGACGC CCAACCTGCC ATCACGAGAT
 5751 TTGATTCCA CCGCCGCTT CTATGAAAGG TTGGGCTTCG GAATGTTT
 45 5801 CCGGGACGCC GGCTGGATGA TCCTCCAGCG CGGGGATCTC ATGCTGGAGT
 5851 TCTTCGCCCA CCCCCAACTTG TTTATTGCAG CTTATAATGG TTACAAATAA
 5901 AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTT CACTGCATTC
 5951 TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGTATAC
 6001 CGTCGACCTC TAGCTAGAGC TTGGCGTAAT CATGGTCATA GCTGTTCTC
 6051 GTGTGAAATT GTTATCCGCT CACAATTCCA CACAACATAC GAGCCGGAAAG

5 6101 CATAAAAGTGT AAAGCCTGGG GTGCCTAATG AGTGAGCTAA CTCACATTAA
 6151 TTGCGTTCGCG CTCACTGCC C GCTTTCCAGT CGGGAAACCT GTCGTGCAG
 6201 CTGCAATTAAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG
 6251 GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCTGCTCGC
 6301 TGCAGCGAGC GGTATCAGCT CACTCAAAGG CGGTAAATACG GTTATCCACA
 6351 GAATCAGGGG ATAACCGCAGG AAAGAACATG TGAGCAAAAG GCCAGCAAAA
 6401 GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCCTTTTC CATAGGCTCC
 6451 GCCCCCCTGA CGAGCATTAC AAAAAATCGAC GCTCAAGTCA GAGGTGGCGA
 6501 AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT
 6551 CGTGCCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCGCCCT
 6601 TTCTCCCTTC GGGAAAGCGTG GCGCTTCTC AATGCTCACG CTGTAGGTAT
 6651 CTCAGTTCGG TGTAGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC
 6701 CCCCCTTCAG CCCGACCGCT GCGCCTTATC CGGTAACATAT CGTCTTGAGT
 6751 CCAACCCGGT AAGACACGAC TTATGCCAC TGGCAGCAGC CACTGGTAAC
 6801 AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG
 6851 GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTTGGT ATCTGCGCTC
 6901 TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC
 6951 AAACAAACCA CCGCTGGTAG CGGTGGTTT TTTGTTGCA AGCAGCAGAT
 7001 TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCCTTGATC TTTTCTACGG
 7051 GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
 7101 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAATT AAAAATGAAG
 7151 TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC
 7201 AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTCGTTCA
 7251 TCCATAGTTG CCTGACTCCC CGTCGTGTAG ATAACCTACGA TACGGGAGGG
 7301 CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCCGAGAC CCACGCTCAC
 7351 CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCGAGCGC
 7401 AGAAGTGGTC CTGCAACTTT ATCCGCTCC ATCCAGTCTA TTAATTGTTG
 7451 CGGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTG CGAACGTTG
 7501 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT
 7551 TCATTCAAGCT CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCAT
 7601 GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCCTCGATC GTTGTAGAA
 7651 GTAAGTTGGC CGCAGTGTAA TCACTCATGG TTATGGCAGC ACTGCATAAT
 7701 TCTCTTACTG TCATGCCATC CGTAAGATGC TTTCTGTGA CTGGTGTAGA
 7751 CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGGGACCG AGTTGCTCTT
 7801 GCGCGGGTC AATACGGGAT AATACCGCC CACATAGCAG AACTTAAAAA
 7851 GTGCTCATCA TTGGAAAACG TTCTCGGGG CGAAAACCT CAAGGATCTT
 7901 ACCGCTGTGAG AGATCCAGTT CGATGTAACC CACTCGTGA CCCAAGTGT
 7951 CTTCAGCATC TTTTACTTTC ACCAGCGTTT CTGGGTGAGC AAAAACAGGA
 8001 AGGCAGGAAATG CCGCAAAAAA GGGATAAGG GCGACACGGG AATGTTGAAT

5 8051 ACTCATACTC TTCCCTTTTC AATATTATTG AAGCATTAT CAGGGTTATT
 8101 GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA
 8151 GGGGTTCCGC GCACATTCC CCGAAAAGTG CCACCTGACG TC

10 Tabelle 10: Nukleotidsequenz des rekombinanten Plasmids pIND-H-Semal-EE (SEQ ID NO.: 37)

15 1 AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT
 51 TGTTCTCGTT AGCTCGATGG ACAAGTGCAT TGTTCTCTTG CTGAAAGCTC
 101 GATGGACAAG TGCAATTGTT TCTTGCTGAA AGCTCGATGG ACAAGTGCAT
 151 TGTTCTCTTG CTGAAAGCTC AGTACCCGGG AGTACCCCTCG ACCGCCGGAG
 201 TATAAAATAGA GGGCCTTCGT CTACGGAGCG ACAATTCAAT TCAAACAAGC
 251 AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT
 301 GAACAAGCTA AACAACTCTGC AGTAAAGTGC AAGTAAAGT GAATCAATTA
 351 AAAGTAACCA GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA
 401 GAAGTAATTA TTGAATACAA GAAGAGAACT CTGAATACTT TCAAACAAGTT
 451 ACCGAGAAAG AAGAACTCAC ACACAGCTAG CGTTTAAACT TAAGCTTGGT
 501 ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGgaattcgg ctggggatga
 551 cgcctccctcc gcccggacgt gcccggccca gcgcacccgcg cgcggcgatgc
 601 cctggcccgcc cggctcggtt gggcttccg ctggcgctgc ggctgctgct
 651 gctgctctgg cggccggccg cctccggccca gggccaccta aggagcggac
 701 cccgcacatctt cggcgtctgg aaaggccatg tagggcagga cccgggtggac
 751 tttggccaga ctgagccgca cacgggtctt ttccacgagc caggcagctc
 801 ctctgtgtgg gtgggaggac gtggcaaggt ctacctcttt gacttccccc
 851 agggcaagaa cgcacatctgtg cgcacgggtga atatcgctc cacaagggg
 901 tcctgtctgg ataagcggga ctgcgagaac tacatcaactc tcctggagag
 951 gcggagtgag gggctgctgg cctgtggcac caacggccgg caccggcagct
 1001 gctggaaacct ggtaatggc actgtgtgtc cacttggcga gatgagaggg
 1051 tacggccccc tcaagccggga cgagaactcc ctgggtctgt ttgaagggga
 1101 cgaggtgtat tccaccatcc ggaaggcagga atacaatggg aagatccctc
 1151 ggttccggcc catccggggc gagagtggac tgtacaccag tgatactgtc
 1201 atgcagaacc cacagttcat caaagccacc atcgtgcacc aagaccaggc
 1251 ttacgatgac aagatctact acttcttccg agaggacaat cctgacaaga
 1301 atcctgagggc tcctctcaat gtgtcccgtg tggcccaagtt gtgcagggggg
 1351 gaccagggtg gggaaagtgc actgtcagtc tccaaagtggc acactttct
 1401 gaaagccatg ctggtatgca gtgtatgtc caccaacaag aacttcaaca
 1451 ggctgcaaga cgtctccctg ctccctgacc ccagcggcca gtggagggac
 1501 accaggggtct atgggtttt ctccaaacccc tggaaactact cagccgtctg
 1551 tgtgtattcc ctcggtgaca ttgacaaggt cttccgtacc tcctcactca

1601 agggctacca ctcaagcatt cccaaacccgc ggcctggcaa gtgcctccca
 1651 gaccagcagc cgataccac agagacccatc cagggtggctg accgtcacc
 5 1701 agagggtggc cagagggtgg agcccatggg gcctctgaag acgcattgt
 1751 tccactctaa ataccactac cagaaagtgg ccgtcaccc catgcaagcc
 1801 agccacgggg agaccttca tgtgcttac ctaactacag acaggggcac
 1851 tatccacaag gtggtggaaac cgggggagca ggagcacagc ttcgccttca
 10 1901 acatcatgaa gatccagccc ttccgcgc cggctgcacat ccagaccatg
 1951 tcgctggatg ctgagcggag gaagctgtat gtgagctccc agtgggaggt
 2001 gagccaggtg cccctggacc tgtgtgaggt ctatggccgg ggctgcccacg
 2051 gttgcctcat gtcccagac ccctactgcg gctgggacca gggccgctgc
 15 2101 atctccatct acagctccga acggctcgtg ctgcaatcca ttaatccagc
 2151 cgagccacac aaggagtgtc ccaaccccaa accagacaag gccccactgc
 2201 agaaggtttc cctggccca aactctcgct actacctgag ctgccccatg
 2251 gaatcccggc acgccaccta ctcatggcgc cacaaggaga acgtggagca
 2301 gagctgcgaa cctggtcacc agagcccaa ctgcattctg ttcatcgaga
 2351 acctcacgac gcagcagttac ggccactact tctgcgagggc ccaggaggc
 2401 tcctacttcc gcgagctca gcactggcag ctgctgccc aggacggcat
 2451 catggccgag cacctgctgg gtcatgcctg tgccctggct gcctccctct
 25 2501 ggctgggggt gctgcccaca ctcactcttgc gcttgcgttgc ccacgtgaag
 2551 cttGGGGCCCG TTTAAACCCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
 2601 CCAGCCATCT GTTGTGAGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG
 2651 GTGCCACTCC CACTGTCCCTT TCCTAATAAA ATGAGGAAT TGCATCGCAT
 30 2701 TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGGACAG
 2751 CAAGGGGGAG GATTGGAAG ACAATAGCAG GCATGCTGGG GATGCGGTGG
 2801 GCTCTATGGC TTCTGAGGCG GAAAGAACCA GCTGGGGCTC TAGGGGGTAT
 2851 CCCCCACGCGC CCTGTAGCGG CGCATTAAGC GCGGGGGTG TGGTGGTTAC
 2901 GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC GCTCCTTTCG
 2951 CTTCTTCCC TTCTTCTC GCCACGTTCG CCGGCTTCC CCGTCAAGCT
 3001 CTAAATCGGG GCATCCCTT AGGGTCCGA TTAGTGCTT TACGGCACCT
 40 3051 CGACCCAAA AAACCTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC
 3101 CCTGATAGAC GGTTTTCGC CCTTGACGT TGGAGTCCAC GTTCTTTAAT
 3151 AGTGGACTCT TGTTCAAAC TGAACAAACA CTCAACCCCTA TCTCGGTCTA
 3201 TTCTTTGAT TTATAAGGG A TTTGGGGAT TCGGGCTTAT TGGTTAAAAA
 45 3251 ATGAGCTGAT TTAACAAAAA TTTAACGCGA ATTAATTCTG TGGAAATGTGT
 3301 GTCAGTTAGG GTGTGAAAG TCCCCAGGCT CCCCCAGGCAG GCAGAAAGTAT
 3351 GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG
 3401 GCTCCCCAGC AGGCAGAAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
 50 3451 ACCATAGTCC CGCCCCCTAAC TCCGCCCCATC CCGCCCCCTAA CTCCGCCCCAG
 3501 TTCCGCCCAT TCTCCGCCCA ATGGCTGACT AATTTTTTTT ATTTATGCAG

3551 AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC
 3601 TTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCCCGGAG CTTGTATATC
 3651 CATTTCGGA TCTGATCAAG AGACAGGATG AGGATCGTT CGCATGATTG
 3701 AACAAAGATGG ATTGCACGCA GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA
 3751 TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT
 3801 GTTCCGGCTG TCAGCGCAGG GGCGCCCGGT TCTTTTGTC AAGACCGACC
 3851 TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG
 3901 CTGGCCACGA CGGGCGTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA
 3951 AGCGGGAAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC
 4001 TGTCACTCA CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA
 4051 ATGCGGGGGC TGCAATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA
 4101 AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG
 4151 TCGATCAGGA TGATCTGGAC GAAGAGCATIC AGGGGCTCGC GCCAGCCGAA
 4201 CTGTTGCCA GGCTCAAGGC GCGCATGCC GACGGGGAGG ATCTCGTCGT
 4251 GACCCATGGC GATGGCTGCT TGCGAATAT CATGGTGGAA AATGGCCGCT
 4301 TTTCTGGATT CATCGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG
 4351 GACATAGCGT TGGCTACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG
 4401 GGCTGACCGC TTCCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTCGCAGC
 4451 GCATGCCCTT CTATGCCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG
 4501 GGTCGAAAT GACCGACCAA GCGACGCCA ACCTGCCATC ACCAGATTTC
 4551 GATTCCACCG CCGCCTTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTCCG
 4601 GGACGCGGGC TGGATGATCC TCCAGCGCG GGATCTCATG CTGGAGTTCT
 4651 TCGCCCACCC CAACTTGT TTGCACTGTT ATAATGGTTA CAAATAAAGC
 4701 AATAGCATCA CAAATTCAC AAATAAAGCA TTTTTTCAC TGCAATTCTAG
 4751 TTGTGGTTG TCCAAACTCA TCAATGTATC TTATCATGTC TGATAACCGT
 4801 CGACCTCTAG CTAGAGCTTG GCGTAATCAT GGTCACTAGCT GTTCCCTGTG
 4851 TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT
 4901 AAAGTGTAAA GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
 4951 CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG
 5001 CATTAAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTGTC GTATTGGCG
 5051 CTCTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
 5101 GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
 5151 TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
 5201 CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTCCAT AGGCTCCGCC
 5251 CCCCTGACGA GCATCACAAA ATCGACGCT CAAGTCAGAG GTGGCGAAAC
 5301 CCGACAGGAC TATAAAGATA CCAGGGCTTT CCCCCCTGGAA GCTCCCTCGT
 5351 GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC
 5401 TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
 5451 AGTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC

5 5501 CGTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
 5551 ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
 5 5601 ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG
 5651 GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC
 5701 TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGAAA
 5751 CAAACCCACCG CTGGTAGCGG TGGTTTTTT GTTGCAAGC AGCAGATTAC
 10 5801 GCGCAGAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT
 5851 CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTT GGTCATGAGA
 5901 TTATCAAAAA GGATCTCAC CTAGATCCTT TAAATTAAA AATGAAGTTT
 5951 TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
 15 6001 GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCATT TCGTTCATCC
 6051 ATAGTTGCCCT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT
 6101 ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG
 6151 CTCCAGATT ATCAGCAATA AACCAAGCCAG CCGGAAGGGC CGAGCGCAGA
 20 6201 AGTGGTCCTG CAACTTATC CGCCTCCATC CAGTCTATT ATTGTTGCCG
 6251 GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTGCGC AACGTTGTTG
 6301 CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTGG TATGGCTTCA
 6351 TTCAGCTCCG GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCCATGTT
 25 6401 GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGT GTCAAGAAGTA
 6451 AGTTGGCCGC AGTGTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT
 6501 CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTCAGT GTGAGTACTC
 6551 AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTGCC
 30 6601 CGGCGTCAAT ACAGGATAAT ACCGGGCCAC ATAGCAGAAC TTTAAAAGTG
 6651 CTCATCATTG GAAAACGTT TTCGGGGCGA AACTCTCAA GGATCTTACC
 6701 GCTGTTGAGA TCCAGTCGA TGTAACCCAC TCGTGCACCC AACTGATCTT
 6751 CAGCATTT TACTTCACC AGCGTTCTG GGTGAGCAAA AACAGGAAGG
 6801 CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGAAAT GTGAAACT
 6851 CATACTCTTC CTTTTCAAT ATTATTGAAG CATTATCAG GGTTATTGTC
 6901 TCATGAGCGG ATACATATT GAATGTATT AGAAAAATAA ACAAAATAGGG
 6951 GTTCCGCGCA CATTCCCCG AAAAGTGCCA CCTGACGTCG ACGGATCGGG
 40

Tabelle 11: Nukleotidsequenz des rekombinanten Plasmids pIND-H-Semal-EA (SEQ ID NO.: 38)

45

50

1 AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT
 51 TGTTCTCGTT AGCTCGATGG ACAAGTGCAT TGTTCTCTTG CTGAAAGCTC
 101 GATGGACAAG TGCATTGTT TCCTGCTGAA AGCTCGATGG ACAAGTGCAT
 151 TGTTCTCTTG CTGAAAGCTC AGTACCCGGG AGTACCCCTCG ACCGCCGGAG
 201 TATAAAATAGA GGCGCTTCGT CTACGGAGCG ACAATTCAAT TCAAACAAGC

55

251 AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT
 301 GAACAAGCTA AACAAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATT
 5 351 AAAGTAACCA GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA
 401 GAAGTAATTA TTGAATACAA GAAGAGAACT CTGAATACTT TCAACAAAGTT
 451 ACCGAGAAG AAGAACTCAC ACACAGCTAG CGTTTAAACT TAAGCTTGGT
 501 ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGaaattcgg cttggatga
 10 551 cgcctcctcc gccggacgt gcccggccca gcgcacccgac cgccggcgac
 601 cctggcccgcccg cggctcggtt gggcttcccg ctgcggctgc ggctgctgct
 651 gctgctctgg gcggccggcccg cctccggccca gggccaccta aggagcggac
 701 cccgcattt cggctctgg aaaggccatg tagggcagga cccgggtggac
 15 751 tttggccaga ctgagccgca cacggctt ttccacgagc caggcagctc
 801 ctctgtgtgg gtggaggac gtggcaagggt ctacctctt gacttccccg
 851 agggcaagaa cgcattctgtg cgcacggta atatcggtc cacaagggg
 901 tcctgtctgg ataagcgga ctgcgagaac tacatcactc tcctggagag
 951 gcggagttag gggctctgg cctgtggcac caacggcccg caccggcgt
 1001 gctggAACCT ggtgaatggc actgtgtgc cacttggcga gatgagaggc
 1051 tacggcccttc tcaagccggc cgagaactcc ctggttctgt ttgaagggg
 1101 cgagggtgtat tccaccatcc ggaaggcagga atacaatggg aagatccctc
 1151 gttccggccg catccggggc gagagttagc tgtacaccag tgatactgtc
 1201 atgcagaacc cacagttcat caaagccacc atcgtgcacc aagaccaggc
 1251 ttacgatgac aagatctact acttcttcccg agaggacaat cctgacaaga
 1301 atcctgaggc tcctctcaat gtgtccctgt tggccctgtt gtgcagggg
 1351 gaccagggtg gggaaagttc actgtcagtc tccaaatggc acactttct
 1401 gaaaggccatg ctggatgca gtgatgtgc cacaacaag aacttcaaca
 1451 ggctgcaaga cgtcttcctg ctccctgacc ccagggccca gtggagggg
 1501 accagggtct atgggtttt ctccaacccc tggaaactact cagccgtctg
 1551 tgtgtattcc ctccggatca ttgacaagggt ctccctgttccctactca
 1601 agggctacca ctcacccctt cccaaacccgc ggcctggcaa gtgcctccca
 1651 gaccagcagc cgatacccac agagacccctc cagggtggctg accgtcacc
 1701 agagggtggc cagagggtgg agcccatggg gcctctgaag acgcccattgt
 1751 tccactctaa ataccactac cagaaatggg ccgttccaccg catgcaagcc
 1801 agccacgggg agacccctca tggctttac ctaactacag acaggggcac
 1851 tatccacaag gtgggtggac cggggggagca ggagcacagc ttccgccttca
 1901 acatcatgga gatccagccc ttccggcccg cggctgccc ccagaccatg
 1951 tcgctggatg ctgagccggag gaagctgtat gtggactccc agtggggaggt
 2001 gagccagggtg cccctggacc tggatgggt ctatggccgg ggctgcccac
 2051 gttgcctcat gtcccgagac ccctactgac gctggacca gggccgctgc
 2101 atctccatct acagctccga acggtcagtg ctgcaatcca ttaatccac
 2151 cgagccacac aaggagtgtc ccaacccaa accagacaag gccccactgc

55

2201 agaaggtttc cctggcccca aactctcgct actacctgag ctgccccatg
 2251 gaatcccgcc acgccaccta ctcatggcgc cacaaggaga acgtggagca
 5 2301 gagctgcgaa cctggtcacc agagcccaa ctgcacccctg ttcatcgaga
 2351 acctcacggc gcagcagttac ggcactact tctgcgaggc ccaggagggc
 2401 tcctacttcc gcgaggctca gcactggcag ctgctcccc aggacggcat
 2451 catggccgag cacctgctgg gtcatgcctg tgccctggct gcctccctct
 10 2501 ggctgggggt gctgcccaca ctcactcttgc gcttgcgtt ccacgtgaag
 2551 cttGGGGCCCG AACAAAAACT CATCTCAGAA GAGGATCTGA ATAGCGCCGT
 2601 CGACCATCAT CATCATCATC ATTGAGTTA TCCAGCACAG TGGCGGCCGC
 2651 TCGAGTCTAG AGGGCCCGTT TAAACCCGCT GATCAGCCTC GACTGTGCCT
 15 2701 TCTAGTTGCC AGCCATCTGT TGTTGCCCTC TCCCCCGTGC CTTCCCTTGAC
 2751 CCTGGAAGGT GCCACTCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
 2801 CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG
 2851 CAGGACAGCA AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGA
 2901 TGCGGTGGGC TCTATGGCTT CTGAGGCGGA AAGAACCGAGC TGGGGCTCTA
 2951 GGGGGTATCC CCACGGCCCG TGTAGCGGGCG CATTAGCGC GGCGGGTGTG
 3001 GTGGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCCCG TAGGCCCGC
 3051 TCCCTTCGCT TTCTCCCTT CTTTCTCGC CACGTTGCCCG GGCTTCCCCC
 3101 GTCAAGCTCT AAATGGGGC ATCCCTTTAG GGTTCCGATT TAGTGCTTTA
 3151 CGGCACCTCG ACCCCAAAAAA ACTTGATTAG GGTGATGGTT CACGTAGTGG
 3201 GCCATCGCCCG TGATAGACGG TTTTCGCCCTT TTTGACGTTG GAGTCCACGT
 3251 TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCATAC
 3301 TCGGTCTATT CTTTGATTT ATAAGGGATT TTGGGGATTT CGGCCTATTG
 3351 GTAAAAAAAT GAGCTGATTT AACAAAAAATT TAACGCGAAT TAATTCTGTG
 3401 GAATGTGTGT CAGTTAGGGT GTGGAAAGTC CCCAGGCTCC CCAGGCAGGC
 3451 AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA
 3501 GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT
 3551 AGTCAGCAAC CATACTCCCG CCCCTAACTC CGCCCATCCC GCCCCCTAACT
 3601 CCGCCCAGTT CCGCCCCATTC TCCGCCCAT GGCTGACTAA TTTTTTTTAT
 3651 TTATGCAGAG GCCGAGGCCG CCTCTGCCTC TGAGCTATTC CAGAAGTAGT
 3701 GAGGAGGCTT TTTGGAGGC CTAGGCTTTT GCAAAAAGCT CCCGGGAGCT
 3751 TGATATATCCA TTTTCGGATC TGATCAAGAG ACAGGATGAG GATCGTTCG
 3801 CATGATTGAA CAAGATGGAT TGCAAGCAGG TTCTCCGGCC GCTTGGGTGG
 3851 AGAGGCTATT CGGCTATGAC TGGCACAAC AGACAATCGG CTGCTCTGAT
 3901 GCCGCCGTGT TCCGGCTGTC AGCCGAGGGG CGCCCGGTTC TTTTTGTCAA
 3951 GACCGACCTG TCCGGTGCCTC TGAATGAACG GCAGGACGAG GCAGCGCGC
 4001 TATCGTGGCT GGCCACGACG GGCGTTCCCTT GCGCAGCTGT GCTCGACGTT
 4051 GTCACTGAAG CGGGAAAGGGG CTGGCTGCTA TTGGGCGAAG TGCCGGGGCA
 4101 GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA TCCATCATGG

4151 CTGATGCAAT GCGGGGGCTG CATA CGCTTG ATCCGGCTAC CTGCCCATTC
 4201 GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC
 4251 CGGTCTTGTG GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC
 4301 CAGCCGAACT GTTCCGCCAGG CTCAAGGCAGC GCATGCCCGA CGGCGAGGAT
 4351 CTCGTCGTGA CCCATGGCGA TGCCTGCTTG CGAATATCA TGGTGGAAAA
 4401 TGGCCGCTT TCTGGATTCA TCGACTGTGG CGGCTGGGT GTGGCGGACC
 4451 GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA AGAGCTTGGC
 4501 GGCAGATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCAG
 4551 TTCGCAGCGC ATCGCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG
 4601 GACTCTGGGG TTCGAAATGA CCGACCAAGC GACGCCAAC CTGCCATCAC
 4651 GAGATTCGA TTCCACCGCC GCCTTCTATG AAAGGTTGGG CTCGGAATC
 4701 GTTTCCGGG ACGCCGGCTG GATGATCCTC CAGCGCGGGG ATCTCATGCT
 4751 GGAGTTCTTC GCCCACCCCA ACTTGTATG TGAGCTTAT AATGGTTACA
 4801 AATAAAGCAA TAGCATCACA AATTCACAAA ATAAAGCATT TTTTCACTG
 4851 CATTCTAGTT GTGGTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG
 4901 TATACCGCTCG ACCTCTAGCT AGAGCTTGGG GTAATCATGG TCATAGCTGT
 4951 TTCCGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA CATA CGAGCC
 5001 GGAAGCATAA AGTGTAAAGC CTGGGGTGCCTA ATATGAGTGA GCTAACTCAC
 5051 ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCGT
 5101 GCCAGCTGCA TTAATGAATC GGCCAACCGCG CGGGGAGAGG CGGTTTGCCT
 5151 ATTGGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTGCT
 5201 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTAA ATACGGTTAT
 5251 CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
 5301 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
 5351 GCTCCGCCCA CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT
 5401 GGC GAAACCC GACAGGACTA TAAAGATACC AGGCGTTCC CCCTGGAAAGC
 5451 TCCCTCGTGC GCTCCCTGT TCCGACCCCTG CCGCTTACCG GATA CCTGTC
 5501 CGCTTCTCTC CCTTCGGGAA CGGTGGCGCT TTCTCAATGC TCACGCTGTA
 5551 GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC
 5601 GAAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
 5651 TGAGTCAAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
 5701 GAAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTG
 5751 AAGTGGTGGC CTAAC TACAGG CTACACTAGA AGGACAGTAT TTGGTATCTG
 5801 CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT
 5851 CCGGCAAACA AACCAACCGCT GGTAGCGGTG GTTTTTTGT TTGCAAGCAG
 5901 CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC
 5951 TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTGG
 6001 TCATGAGATT ATCAAAAGG ATCTTCACCT AGATCCTTT AAATTAAAAA
 6051 TGAAGTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG

6101 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTC
 6151 GTTCATCCAT AGTTGCCCTGA CTCCCCGTG TGAGATAAC TACGATACGG
 5 6201 GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG
 6251 CTCACCGGCT CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG
 6301 AGGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT
 6351 TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTGCGCAA
 10 6401 CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTGGTA
 6451 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
 6501 CCCATGTTGT GCAAAAAAGC GTTACGCTCC TTCGGTCCTC CGATCGTTGT
 6551 CAGAAGTAAG TTGGCCGCAG TGTTATCACT CATGGTTATG GCAGCACTGC
 15 6601 ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT
 6651 GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCCGG GACCGAGTTG
 6701 CTCTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT
 6751 TAAAAGTGC CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG
 20 6801 ATCTTACCGC TGTTGAGATC CAGTCGATG TAACCCACTC GTGCACCCAA
 6851 CTGATCTTCA GCATCTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
 6901 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT
 6951 TGAATACTCA TACTCTTCCT TTTCAATAT TATTGAAGCA TTTATCAGGG
 25 7001 TTATTGTCTC ATGAGCGGAT ACATATTGA ATGTATTTAG AAAAATAAAC
 7051 AAATAGGGGT TCCGCGCACA TTTCCCGAA AAGTGCCACC TGACGTCGAC
 7101 GGATCGGG

30

Tabelle 12: Sequenz des rekombinanten Plasmids pQE30-H-Semal-BH (SEQ ID NO.: 39)

35 1 CTCGAGAAAT CATAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT
 51 AATAGATTCA ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
 101 AGGAGAAATT AACTATGAGA GGATCGCATC ACCATCACCA TCACGGAtcc
 151 ctggttctgt ttgaagggga cgaggtgtat tccaccatcc ggaagcagga
 40 201 atacaatggg aagatccctc gttccgcgg catccggggc gagagtggc
 251 tgtacaccag tgatactgtc atgcagaacc cacagttcat caaagccacc
 301 atcgtgcacc aagaccaggc ttacgatgac aagatctact acttcttccg
 351 agaggacaat cctgacaaga atcctgaggc tcctctcaat gtgtcccgtg
 401 tggcccagtt gtgcaggggg gaccagggtg gggaaagttc actgtcagtc
 451 tccaaagtggaa acactttct gaaagccatg ctggatgtca gtgtatgtc
 501 caccacaaca aacttcaaca ggctgcaaga cgtcttcctg ctccctgacc
 551 ccagcggcca gtggagggac accagggtct atgggtttt ctccaaacccc
 601 tggaaactact cagccgtctg tggatattcc ctggtgaca ttgacaagg
 651 cttccgttacc tcctcactca agggctacca ctcaagcctt cccaaacccgc

55

701 ggcctggcaa gtgcctccca gaccagcagc cgataccac agaAAGCTTA
 751 ATTAGCTGAG CTTGGACTCC TGTTGATAGA TCCAGTAATG ACCTCAGAAC
 5 801 TCCATCTGGA TTTGTTAGA ACGCTCGGTT GCGCAGGGC GTTTTTATT
 851 GGTGAGAATC CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA
 901 AAATGGAGAA AAAATCACT GGATATACCA CCGTTGATAT ATCCCAATGG
 951 CATCGTAAAG AACATTTGA GGCATTCAG TCAGTTGCTC AATGTACCTA
 1001 TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTAAAG ACCGTAAAGA
 1051 AAAATAAGCA CAAGTTTAT CCGGCCTTA TTCACATTCT TGCCCGCCTG
 1101 ATGAATGCTC ATCCGAATT TCGTATGGCA ATGAAAGACG GTGAGCTGGT
 1151 GATATGGGAT AGTGGTCACC CTTGTTACAC CGTTTCCAT GAGCAAACCTG
 1201 AACGTTTC ATCGCTCTGG AGTGAATACC ACGACGATT CCGGCAGTTT
 1251 CTACACATAT ATTGCAAGA TGTGGCGTGT TACGGTGAAA ACCTGGCCTA
 1301 TTTCCCTAAA GGGTTTATTG AGAATATGTT TTTCGTCTCA GCCAATCCCT
 1351 GGGTGAGTTT CACCAGTTT GATTAAACG TGGCAATAT GGACAACCTTC
 20 1401 TTCGCCCCCG TTTTACCCAT GGGCAAATAT TATAGCAAG GCGACAAGGT
 1451 GCTGATGCCG CTGGCGATTC AGGTTCATCA TGCCGCTGTG GATGGCTTCC
 1501 ATGTCGGCAG AATGCTTAAT GAATTACAAAC AGTACTGCGA TGAGTGGCAG
 1551 GGCGGGGCGT AATTTTTTA AGGCAGTTAT TGGTGCCCTT AAACGCCTGG
 25 1601 GGTAATGACT CTCTAGCTTG AGGCATCAA TAAAACGAAA GGCTCAGTCG
 1651 AAAGACTGGG CCTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT
 1701 GAGTAGGACA AATCCGCCGC TCTAGAGCTG CCTCGCGCGT TTCGGTGATG
 1751 ACGGTGAAAA CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT
 30 1801 CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG
 1851 TGGTGGCGGG TGTGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG
 1901 GAGTGTATAC TGGCTTAACT ATGCGGCATC AGAGCAGATT GTACTGAGAG
 1951 TGCACCATAT GCGGTGTGAA ATACCGACA GATGCGTAAG GAGAAAATAC
 2001 CGCATCAGGC GCTCTTCCGC TTCTCGCTC ACTGACTCGC TGCGCTCGGT
 2051 CTGTCGGCTG CGGGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT
 2101 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
 40 2151 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTCCA
 2201 TAGGCTCCGC CCCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA
 2251 GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA
 2301 AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCGCTTA CGGGATAACCT
 45 2351 GTCCGCTTT CTCCCTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT
 2401 GTAGGTATCT CAGTTCGGTG TAGGTCGTTG GCTCCAAGCT GGGCTGTGTG
 2451 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG
 2501 TCTTGAGTCC AACCCGGTAA GACACGACTT ATGCCACTG GCAGCAGCCA
 50 2551 CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC
 2601 TTGAAGTGGT GGCTAACTA CGGCTACACT AGAAGGACAG TATTGGTAT

2651 CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
 2701 GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTT TTGTTGCAAG
 2751 CAGCAGATT CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
 2801 TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATT
 2851 TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA
 2901 AAATGAAGTT TAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA
 2951 CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
 3001 TTCGTTCATC CATAGCTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
 3051 CGGGAGGGCT TACCATCTGG CCCCCAGTGC GCAATGATAC CGCGAGACCC
 3101 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCCAGCCA GCCGGAAGGG
 3151 CCGAGCCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT
 3201 AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTGCG
 3251 CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTG
 3301 GTATGGCTTC ATTCAAGCTCC GGTCCCCAAC GATCAAGGCG AGTTACATGA
 3351 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT
 3401 TGTCAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC
 3451 TGCATAATTTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT
 3501 GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCAGCCGAG
 3551 TTGCTCTTGC CGCGCGTCAA TACGGGATAA TACCGGCCA CATAGCAGAA
 3601 CTTTAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA
 3651 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
 3701 CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTCT GGGTGAGCAA
 3751 AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA
 3801 TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA GCATTTATCA
 3851 GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
 3901 AACAAATAGG GGTTCCGCAC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
 3951 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCCTATCAC
 4001 GAGGCCCTT CGTCTTCAC

40

Tabelle 13: Sequenz des rekombinanten Plasmids pQE31-H-Semal-SH (SEQ ID NO.: 40)

1 CTCGAGAAAT CATAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT
 51 AATAGATTCA ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
 101 AGGAGAAATT AACTATGAGA GGATCGCATC ACCATCACCA TCACACGGAT
 151 CCGCATGC_ga gctcccagt_g ggaggtgagc caggtggccc tggacctgt_g
 201 tgaggcttat ggcggggct gcccacgggtt cctcatgtcc cgagacccct
 251 actgcggctg ggaccagggc cgctgcacatct ccatctacag ctccgaacgg
 301 tcagtgctgc aatccattaa tccagccgag ccacacaagg agtgtcccaa

55

5	351	ccccaacca gacaaggccc cactgcagaa ggttccctg gccccaaact
	401	ctcgctacta cctgagctgc cccatggaaat cccgcacgc cacctactca
	451	tggcgccaca aggagaacgt ggagcagagc tgcgaacctg gtcaccagag
	501	ccccaaactgc atccctgttca tcgagaacct cacggcgcag cagtacggcc
	551	actacttctg cgaggcccag gagggctcct acttccgcga ggctcagcac
10	601	tggcagctgc tgcccgagga cggcatcatg gccgagcacc tgctgggtca
	651	tgcctgtgcc ctggctgcct ccctctggct ggggggtgctg cccacactca
	701	ctcttggctt gctggtccac gtgaagcttA ATTAGCTGAG CTTGGACTCC
	751	TGTTGATAGA TCCAGTAATG ACCTCAGAAC TCCATCTGGA TTTGTTCA
15	801	ACGCTCGGTT GCCGCCGGGC GTTTTTATT GGTGAGAATC CAAGCTAGCT
	851	TGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT
	901	GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTGA
	951	GGCATTCAG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG
20	1001	ATATTACGGC CTTTTAAAG ACCGTAAAGA AAAATAAGCA CAAGTTTAT
	1051	CCGGCCTTTA TTCACATTCT TGCCCGCTG ATGAATGCTC ATCCGGAATT
	1101	TCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC
	1151	CTTGTACAC CGTTTCCAT GAGCAAACGT AAACGTTTC ATCGCTCTGG
	1201	AGTGAATACC ACACGATTT CCGGCAGTTT CTACACATAT ATTGCAAGA
25	1251	TGTGGCGTGT TACGGTAAA ACCTGGCTA TTTCCCTAAA GGGTTTATTG
	1301	AGAATATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTT
	1351	GATTAAACG TGCCCAATAT GGACAACCTC TTCGCCCGGG TTTTCACCAT
30	1401	GGGAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC
	1451	AGGTTCATCA TGCGTCTGT GATGGCTTCC ATGTCGGCAG AATGTTAAT
	1501	GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGCGT AATTTTTTA
	1551	AGGCAGTTT TGGTGCCCTT AAACGCCCTGG GGTAATGACT CTCTAGCTTG
35	1601	AGGCATCAAA TAAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTCGTTT
	1651	TATCTGTTGT TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGC
	1701	TCTAGAGCTG CCTCGCGCGT TTCGGTGTG ACGGTGAAAA CCTCTGACAC
	1751	ATGCGACTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG
40	1801	CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG TGTCGGGCG
	1851	CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAAC
	1901	ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT GCGGTGTGAA
	1951	ATACCGCACA GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTCCGC
45	2001	TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CTGTCGGCTG CGGCGAGCGG
	2051	TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT
	2101	AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAGG CCAGGAACCG
	2151	TAAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC CCCCTGACG
50	2201	AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
	2251	CTATAAAGAT ACCAGGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC

5	2301	TGTTCCGACC CTGCCGCTTA CGGGATACCT GTCCGCCTTT CTCCCTTCGG
	2351	GAAGCGTGGC GCTTTCTCAA TGCTCACGCT GTAGGTATCT CAGTTCGGTG
	2401	TAGGTCGTTG CACTCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC
	2451	CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA
	2501	GACACGACTT ATGCCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
	2551	GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA
10	2601	CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
	2651	TTACCTTCGG AAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAAACCACC
	2701	GCTGGTAGCG GTGGTTTTTG TGTTTGCAG CAGCAGATTA CGCGCAGAAA
	2751	AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC
15	2801	AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
	2851	AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT
	2901	CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
20	2951	GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCTGTCATC CATAGCTGCC
	3001	TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG
	3051	CCCCAGTGC GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
	3101	TATCAGCAAT AAACCAGCCA GCCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
25	3151	GCAACTTTAT CCCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG
	3201	AGTAAGTAGT TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTA
	3251	CAGGCATCGT GGTGTCACGC TCGTCGTTG GTATGGCTTC ATTCAAGCTCC
	3301	GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAAA
30	3351	AGCGGTTAGC TCCTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG
	3401	CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTTC TCTTACTGTC
	3451	ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
	3501	ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGCGTCAA
35	3551	TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT
	3601	GGAAAACGTT CTTCGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG
	3651	ATCCAGTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
40	3701	TTACTTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
	3751	GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
	3801	CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG
	3851	GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC
45	3901	ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTATCAT
	3951	GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT CGTCTTCAC

50 Tabelle 14: (Partielle) Nukleotidsequenz des humanen Semaphorin L Gens.
(8888 Nukleotide) (SEQ ID NO.: 41):

5 GAGCCGCACACGGTGCTTCCACGAGCCAGGCAGCTCCTCTGTGTGGGTGGGAGGACGT
 GGCAAGGTCTACCTCTTGACTTCCCCGAGGGCAAGAACGCATCTGTGCCACGGTGAGC
 CTCTCTCTTCCCCAACACCCCCCTACCCCTTATCTCCCCTCTGGCCCTGCCAAGGGT
 10 CCTCAGGGAAATCCGAGGGAGCTGGCTCTTCTAAACTGCCACCTCCGTATCCTA
 TAAATGGCTCTGGGGAGGCTCCCTAAAGGTAGTCCAGATTGGAGTGGGGAGCTGGGG
 GGTGTGGAGAAAAACAGGAGCTAATGGCCTGGCCAGCTGGCAGCGCTGCGGAAAG
 CCCAGGCTGGAAGCTGGGCCCCAGAGCCATGCCTGGCTTCTGAACCCCTGGGCTCA
 GCTCTGGATATGAGACCTGTTGACCTCAGGTAGATCACTCACCTCTCAGAGCCCCAG
 TTGCTCATCTGTCAAGATGAGAATAATGGTGCTCCTTGGGCTTACCTGAGGCTGTG
 TGGAAGCATTCAGGGTACCTCACCCCTGGCAGATTGAACTAATGCTCTCCCTTCC
 15 CCAGGTGAATATCGGCTCCACAAAGGGTCTGTCTGGATAAGCGGGTGAGGGGGAGG
 GATCTGGAGGGTCTGAGCCACTGGTAAGGGAGAGGACCTGAGGGTCTAAGGAAG
 GAAGCATGGCCCTGCCCCACGAGTCCCAGACTGATGGGAGACGTGGCTCTGTGCTTA
 GGGGATGGCGTCAGCTGCACACACTCTGGGCTGTCCCAGGGTGTGACCTATGCTAAG
 20 CCCCTCTGACACCTTCTCCCTGATCCTGGGGTCTAGTGCCTAGGCTTGCCAGGGCTT
 CCAGCAACCAATTCTCTCCCTTCTCTCTTCCCCTGAGGACTGCGAGAACTACAT
 CACTCTCCCTGGAGAGGCGGAGTGAGGGCTGCTGGCTGTGGCACCAACGCCGGCACCC
 25 CAGCTGCTGGAACCTGGTGAGAAGGCTGCTCCCCATGTGCCTGATCAGCTCACCTCTAC
 TGCCTGGGCTCTGCCCTCATGGTGGAAAGGAGATGGCGAGACTCCATGCTGGCTTG
 CCCTGGGAGGATGGGCTCTGGCCAGAAAAGTGGCGTCATGGGAGGCAGTGGCTGTGG
 GATTATGTGCCATCCAACCCCTGGATCTCCACAGGTGAATGCCACTGTGGTGCCT
 30 TGGCGAGATGAGAGGCTACGCCCTTCAGCCGGACGAGAACTCCCTGGTCTGTTGA
 AGGTTGGGCTGCTCGGAACCTGGCTGGAGCAGGATGGTCAGCTCTTGCCAGTGT
 CCGGAGGAGGGACTTCCAGGAGCTGCCCTACTCATTCTCCCTCCACTGACCCC
 AGGGGACGAGGTGTATTCCACCATCCGGAACAGGAATAACATGGAAAGATCCCTGGT
 35 CCGCCGCATCCGGGGCGAGAGTGAGCTGACACCAGTGATGACTGTCATGCAGAGTGAGTC
 AGGCTCCGGCTGGGCTGAGGGTGGCAAGGGGGTGTGAGCACTTAAGGTGGCAGATGGGA
 TCCTGATGTTCTGGAGGGCTCCCTGAGGGCCCTGGGCAATGCAAGGAAAGCAGGACC
 TTGGTATAGGCCTGAGAAGTTAGGGTGGCTGGGAGCAGAGGAACAGACAAGGTATAGCA
 40 GTGGGATGGGCCAGCCCTTCAGGAACACAAACAGAGGGAGCCCCAGACCCAGTGCAG
 GGTCCCAGGAGCCAAGTTATCCTCTGCTGAGTTACGTTGGAGGCAGCCCCCAACTC
 CCTCCATCAGGGCTCTGCCAATTGAGCAGAAGTGACATAGGGCCCCCAGGGACCTTC
 45 CCCCACCTCCCAGGCATGAAGTCATTGCTCTGGCCATGACATCTTGCTAGGAAGAGG
 GCACAAACAGGTGTGGGTGGAGGTGCAGGGCTAGGGCCCTGGGAGTTGGACCTGAT
 GTTATGAGTCCTATTCCAGATCTGATTGCCATGGTTGTGCAGACCCGAAGGAGGGAGG
 AGAGTGTGCAGGGTGGAAATGGTCTCCGGCAAGCTCCAGCCTTACGCCATTGCT
 50 TCTGTGCCCTGGCAGACCCACAGTCATCAAAGCCACCATCGTGCACCAAGACCAGGCTT
 ACGATGACAAGATCTACTACTCTTCCGAGAGGACAATCCTGACAAGAATCCTGAGGCTC
 CTCTCAATGTGTCCCGTGTGGCCCAGTTGTGCAGGGTGAACACGGCGTGAGGGCTGCTG

5 GCTACGTGTCTGTGCATGAATAGGCCTGAGTGAGGGTGAGTTCTGTGTGTCCGTGTGCAT
 GTAGAAGTTGTGTGGATGTATGAGTGGGTCTGTGTCAAGGACTGTGGGAGCAGCTGTGTG
 TGCATGGAGCATCATGTGTCTGTGTGGTAAAGGTGGCTGAGCCTGTGCACGTATG
 10 ATGGCGTGTGAGCGTGTATGATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
 GTGTGAATGTGCTGTGCCACGTATGTGGGTGCGTGAGTCAGTAAATGTGTGTCTGAGTCC
 GTCTGCTCTGTGGGACCTGGCACTCTCACCTGCCCTGACCCCTGGGCACTGCTGCCCTG
 GGCTCTGGATCAGCCAGGCCTGCTGCAAGGAGTCTCATCTGGAGACCTGCCCTGAGTCCT
 GGGGCACCCCCGGCAGGTCCCTGGCCCTCGCAGCCTGCCCTCCCTCTGGGCCAGGTG
 TTGATATTGCTGGCAGTGGTTCTGGGTGTGTGGGAAAGCTCACTGTCAGTCT
 15 GCCTCTTCTCCCTCACCTTCCAGGGGACCAAGGGTGGGAAAGTTCACTGTCAGTCT
 CCAAGTGGAACACTTTCTGAAAGCCATGCTGGTATGCAGTGATGTCGCCACCAACAAGA
 ACTTCAACAGGCTGCAAGACGTCCTCCCTGCTCCCTGACCCAGCAGGCACTGGAGGGACA
 CCAGGGCTATGGTGTCTCCACCCCTGGTGAGTGCCCTTGTCCTGGGGCCGGG
 20 TGGCATTGGTCAGTGTCCAGTAGGGACAGGAGGCCTGGGCCCTGCTGAGGGCTCCCT
 GGTGTGGCAGGAGCAGGGGCTGCAAGGCTCAAGAGGCTGGCTGTTGCTGGGTGTGGGTG
 GGGGGACAGCCAGTGGCATGTACTGTTGTGAGTGAGTCTGCACTCATGGTG
 TGTGCATGCCCTATATGCACACTCATGACTGCACTTGTGCCTGTGTCCCACACCTGC
 TTGTGCCAGAGTGACACTGGGCCAGGAGGAAGCTGCTGAAGCATCTCGGGGAGCT
 25 GGGTGCTATTACACCTGCTCAGGCACTGCCAGGCCGATAATTACACACTCTTAATCAC
 TCTCATTGATTGAACACACGGCAGGCCAGGCAAGTGTTGGGTGTGTGGGGAGAGTTAGGGA
 TAGAGTGGAGGAAGCCAAGACCCCTGCTCTGTGGCTCCTGGGTGAGTGGGTCCCCCAGGCT
 GGGAAAGGGGTTGGGGCTGGCCCTCTGGGCATCAGCACCCACAGCCTGTGCCAGGG
 30 AGGGCTAGAGAACTGCTCAGCCTATGATGGGTTCTCTGCCCTGGGTGAGTGGAGGTGGATGGG
 TGGGTTCTGGGCCCTCATTGAGCTGGTCTCCATGTCACTGCAGGAACACTCAGCCGTC
 TGTGTGATTCCCTCGGTGACATTGACAAGGTCTCCGTACCTCCTCACTCAAGGGCTAC
 CACTCAAGCCTCCCAACCCGCCCTGGCAAGGTGAGCGTGACACCAGCCGTGGCCAG
 35 GCCCAGCCCTCTGCTCACCTCCCACCCACTGACCTGGCCTGCTCTCCT
 CCCAGTGCCTCCCAGACCAGCAGCCGATAACCCACAGAGACCTCCAGGTGGCTGACCGTC
 ACCCAGAGGTGGCGCAGAGGGTGGAGGCCATGGGCCTCTGAAGACGCCATTGTC
 40 CACTAAATACCAACTACCAAGAAAGTGGCCGTCCACCGCATGCAAGCCAGCCACGGGAGACCT
 TTCATGTGCTTTACCTAACTACAGGTGAGAGGCTACCCGGGACCTCAGTTGCTTTGT
 AAAAACGGGCATGAAAGGTGAAGGAATAATGTAGTTAACATCTGGTTGGATTTACAT
 45 GTGGAGGAATAATTGAGTGACTGGAGTTGTCAGGGGTTAATGTGTGTGGGTGTGGAGA
 GCCAGGGAGAGCTCCCTGGAGGAGGTAGGGCAAGAGGGAAAGGGGGATGGGAGAA
 50 AAGCAAGCACTGGGATTGGAGGGCGAAATCTGGAGAGTCTGAGCAAAGCCAGGTGCACC
 TTTGGTCCAGATGTGACTCAGGGAAAGATGGTAGGAAGAGACGTGGCAAATGAGGA

5 GGAGGGGCCTGAACCACAGGGATACTGGCCTCTGCCAGGCAGAATGAGGGAGTCAGGCC
 TGCCTGCTTTGGATTGTCAGGTGAGAAGAACATTGAGGAGTTGATGGGCACA
 AATTAGGTATGGGAAGGAGTTCAGGGGCAGAACCTTGCATCTCACAGAGGACAGG
 10 GGCAGCTCTCTTCCCTGGAGTAGGCCTGCTGGGAAGCTGGTGGATGCCGTG
 GGAGATGCTCCTGTTCTGAAAGCCACAGGACACGGAGGAGCCAGTCCTGAGTTGGT
 TTGTCGCAGCTCCCAGGCCACAGGACACGGAGGAGCCAGTCCTGAGTTGGT
 TGGGCCATTCAATTCAAGGCCAGGCAGGCCAACCTCAGTTGTCACATTCCCAGTGAT
 15 CTCCCTGGCTGCTTCACCTTGGGACTGTCTGGCTTGGTACCTGTAGGAACTGGA
 ACCCAGCACCATGTTGGCTCTGAAAGCCTTGGGAGAGGAATTCCCACAGGGCAG
 GGCCTGGTCTGATTCCCTGCCTCTTACTCCATTATCATCCGCTACACCTTGGC
 20 CCCCATCCTTGCTGGCTCCAGTACTGGCTGGCACAGCTGGTGTGGTATCCAGGGATGG
 CAGGGCACTGGGAACAGAAGAGAGAGTCACACAGTGGGAACGGAGCAGGAGCTAG
 GACAAGGAAGGCTGGACTTGGCCATGGATTCCCTCTGCAGACTTGGGAAGTGGCAC
 ACTTGAGTGATTAGAGAAGGTTCTCGTCTAAGGGCAGTGGAGGAGGCACCAATTGG
 AGCCTGCATCATCGTATTGGCTAGATTGAAAAATAGAGCTTCTAAGTCCTCTGCAG
 25 AGAATGGGAGGCTCTCACAACTGGAGAAGTATTGGCTTTCTGAGAATTGGCAA
 GGGTATGCTGTTACTGGGCTGGTGGAGGAGTATAGGGCATTATGTCTGTGAAGGCA
 GTGGCTGGGGTGGGGCTTATCAGGCCAAGGAGCATGGCACATCTCAGAGTCACA
 GATGAGGATCACGGATGTGAGAGGAAACATCTAGGCAGGCATCTGACTGCTTT
 TTGGGCAGGTGATGCCCTGGGAAATTGGGAGGGAGAGAGGGAGGTAGGCTATTCT
 30 AGAAACTGGGAGAGCAGGTGAGGTTAGGATGGAGGACCAGGGTCAGGGTCCCCATTGG
 TCCCTAATTGAGAACGGAGAGCATTGGCTAGGAGGCAGCTCGGTTATAAGACC
 TTGGGAACTCTGATTTAGAATCCAAGATCCTTTAGATCTAGGATTTATAAAATTAA
 GATATCCCCTAAAGATCAAATGCAACGTGGAGTCCTGAATTGGATCCTAGAACAGAAG
 GACATTGTGGAAAATAGTGAATCCAATAAAGCTGTAGTTGTTAATAGTAATG
 35 CACCAATGTCAGTGCTAGTTGTGACAATACCGTGGTTATGTAAGATGGTAACATT
 AGGGGAACGGAGGCTGGAGGAGTGGAGCTCTGTACTATCTTGCAACTTTCTGG
 AATCTAAAATTACTCCAAATAAAGGTTAGTAAAGTAATATATTCCCTAAGA
 GTCCAGGAGGCAGGGAGTTGAGAGCAGCTGAGTGGTTGGGTTCTGACAGATTGGT
 40 CCAACTCGGTCTGCTGCTCACAGCTGTGACCTTGAGCAAGTGGCTAGCCTTCT
 GAGCCTGATTTCTTATCTGTGGAGTGGGAAGATGACAGCCACCTCGCAGGGCTGTGGA
 GGGTAAACGAGGTGATGCAGGACAGCAGCCGACTGACCTTGCTGGTGTGGGCTCCT
 GCTCTGTTCTTCCCTGCAAGCTGGGAATGTTGGAGGCCGTATCCAGGGACCCCTGG
 45 CCTCCTGGATGCCCTCTGGATCAGCCTGGAGTCCAGGCTGCCCTTAGGCTCCC
 ACATCTTCCCCAGTCAGCCTCTGCCCTGCCACACCAGTCCTGTGACCTTGCCT
 GAGTTGTGACTTCCCACCCCTCCCCGGCTAGAGGAAAGCTGCTGGCCCTCAGTGGGA
 CTCCCGCCCACTGACCCCTGTCCACCATAACAGACAGGGCACTATCCACAAGGTGGT
 50 GGAACCGGGGAGCAGGAGCACAGCTTGCCTTCAACATCATGGAGATCCAGGCCCTCCG
 CGCGCGGCTGCCATCCAGACCATGTCGCTGGATGCTGAGCGGGTGAGCCTCCCCACT

5 GCGTCCCAGGGCTATGCAGTGACTGCAGCTGAGGACAGGGCTCCTTGCATGTGATTG
 TGTGTTCTTTAAGAGCTCTAGGCCTAGGCCTGGACATTAGGACTGAGTGTTGGGT
 GGGGCCCCGGGCTGACCCAAATCTGCTGCTCCAGAGGAAGCTGTATGTGAGCTCCCA
 10 GTGGGAGGTGAGCCAGGTGCCCTGGACCTGTGTGAGGTCTATGGCGGGGCTGCCACGG
 TTGCCTCATGTCCCAGACCCACTGCGGCTGGGACCAGGGCGCTGCATCTCATCTA
 CAGCTCCGAACGGTACGTTGGCGGGATCCCTCCGTCCTGGACAAAGGTGGCATGGGA
 CAGGGGAGGTGTTGCGGGCTGGAAGAGGTGGCGGTACTGGCCTTCTGTGGGACCT
 15 CCTCTCTACTGGAACGTGACTAGGGTAAGGATATGAGGTCAAGGTCTGAGCCTTGTAT
 CTGCTGATCCTTTCTGCTCCACTCCAGGTCACTGCGCAATCCATTAATCCAGCC
 GAGCCACACAAGGAGTGTCCCAACCCCAACCCAGGTACCTGATCTGGCCTGCTGGCGC
 TGTGGCCAATGAGTGGGTACTGCCCCGCCCCATGGGACAGGATACTGCTGAGAGGGAGCCTCT
 20 CTTTGTCTGTGGGCCCCAGGTACATGGGCAAGGATACTGCTGAGAGGGAGCCTCT
 TGGTGGGATGAGCGAGACGGGAGAAAAAAGGAGGGACGCTGAGGGCTGGGTTCCCCACGTT
 CATTCAAGGCCTGTCCTGGGATCCCAAGTCAGTGGTGGGGAGGACACATCCTCCCCGGAG
 CTCTTGTCCCTCCTCACGGCTGCTCCCCACTGCTCCCCAGACAAGGCCCCACTGCAG
 25 AAGGTTTCCCTGGCCCCAAACTCTCGCTACTACCTGAGCTGCCCCATGAAATCCCAC
 GCCACCTACTCATGGGCCACAAGGAGAACGTTGGAGCAGAGCTGCGAACCTGGTACCCAG
 AGCCCCAACTGCATCCTGTCATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTC
 TGGCAGGGCCCAGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGGGAG
 GACGGCATCATGGCCGAGCACCTGCTGGGTATGCTGCTGCCCCCTGGCCGCTCCCTCTGG
 30 CTGGGGGTGCTGCCACACTCACTCTGGCTTGCTGGTCACTAGGGCCTCCGGAGGCTG
 GGCATGCCCTCAGGCTTCTGCAAGCCAGGGCACTAGAACGTCACACTCAGAGCCGGCTG
 GCCCGGGAGCTCCTGCCACTTCTCCAGGGACAGAATAACCCAGTGGAGGATGC
 CAGGCCTGGAGACGTCCAGCCGAGGCGCTGCTGGGCCCCAGGTGGCGACGGATGGT
 35 AGGGGCTGAGAATGAGGGCACCGACTGTGAAGCTGGGATCGATGACCAAGACTTTAT
 CTTCTGGAAAATATTTCACTCCTCAAACCTGACTAAATGAGCGATGCTCCCAAGCC
 CAAGAGCCCAGGGCTGGGAGTGGTTGGATAGGAGAGCTGGACTCCATCTGACCC
 TGGGGCTGAGGCCTGAGTCCTCTGGACTCTGGTACCCACATTGCTCCTTCCCCCTCC
 40 TCTCTCATGGCTGGTGGCTGGTGTCTGAAAGACCCAGGGCTACCCCTGTCAGCCCT
 GTCCCTCTGCAAGCTCCCTCTGGTCTGGTCCCACAGGACAGCCGCTGATGTTAT
 TGAAGGATTTGCTTCCGGACGGAAGGACGGAAAAAGCTCTGAAAAA
 AAAAAAAA

45 Tabelle 15: Nukleotidsequenz von pMellBacA-HSEMal (6622bp) (SEQ ID NO: 42)

1 GATATCATGG AGATAATTAA AATGATAACC ATCTCGAAA TAAATAAGTA
 50 □
 51 TTTTACTGTT TTCGTAACAG TTTTGTAAATA AAAAAACCTA TAAATATGAA

101 ATTCTTAGTC AACGTTGCC 5 TTGTTTTAT GGTCGTATAC ATTTCTTACA
 151 TCTATGCGGA TCGATGG
 201 cgcacatctcg ccgtctggaa aggccatgt 10 a gggcaggacc gggtggactt
 251 tggccagact gagccgcaca cggtgcttt ccacgagcca ggcagctcct
 301 ctgtgtgggt gggaggacgt ggcaaggct 15 acctcttga ctccccgag
 351 ggcaagaacg catctgtgcg cacggtaat atcggctcca caaaggggtc
 401 ctgtctggat aagcgggact gcgagaacta catca 20 ctactctc ctggagaggc
 451 ggagtgaggg gctgctggcc tggccacca acgccccggca ccccagctgc
 501 tggAACCTGG tgaatggcac 25 tgggtgcca cttggcgaga tgagaggcta
 551 tgcccccttc agcccgacg agaactccct ggttctgtt 30 gaaggggacg
 601 aggtgtattc caccatccgg aagcaggaat acaatggaa gatccctcg
 651 ttccggcgca tccggggcga 35 gagtgagctg tacaccagtg atactgtcat
 701 gcagaaccca cagttcatca aagccaccat cgtgcaccaa gaccaggctt
 751 acgatgacaa gatctactac ttcttccag 40 aggacaatcc tgacaagaat
 801 cctgaggctc ctctcaatgt gtcccgtgt 45 gcccagttgt gcagggggga
 851 ccagggtggg gaaagttcac 50 tgcagtc 1001 caagtggAAC acctttctga
 901 aagccatgct ggtatgcagt gatgctgcca 55 ccaacaagaa cttcaacagg
 951 ctgcaagacg tcttcctgct ccctgacccc 1001 agcggccagt ggagggacac
 1001 cagggtctat ggttttct 1051 ccaaccctg 1051 gaactactca gccgtctgt
 1051 tgtattccct cggtgcatt 1101 gacaaggct 1101 tccgtacctc ctcactcaag
 1101 ggctaccact 1151 caagccttcc 1151 caaccccg 1151 ctttgc 1151 aactacagac
 1151 ccagcagccg 1201 atacccacag 1201 agaccc 1201 ggtggctgac 1201 aggggacta
 1201 aggtggcgca 1251 gagggtggag 1251 cccatggggc 1251 ctctga 1251 tgcaagccag
 1251 cactctaaat 1301 accactacca 1301 gaaagtggcc 1301 gttcaccgca 1301
 1301 ccacggggag 1351 acctttcatg 1351 tgctttac 1351 aactacagac 1351 aggggacta
 1351 tccacaagg 1401 ggtggAAC 1401 ggggagcagg 1401 agcacagctt 1401 cgccttcaac
 1401 atcatggaga 1451 tccagccctt 1451 ccggccgcgc 1451 gctgccatcc 1451 agaccatgtc
 1451 gctggatgct 1501 gagcggagga 1501 agctgtatgt 1501 gagctcccag 1501 tgggaggtga
 1501 gccaggtgcc 1501 cctggacctg 1501 tgtgaggct 1501 atggcgggggg 1501 ctgccacggg

1551 tgcctcatgt cccgagaccc ctactgcggc tgggaccagg gccgctgcat
 1601 ctccatctac agctccgaac ggtcagtgtc gcaatccatt aatccagccg
 1651 agccacacaa ggagtgtccc aaccccaaac cagacaaggc cccactgcag
 1701 aaggtttccc tggcccaaa ctctcgctac tacctgagct gccccatgga
 1751 atcccgcac gccacctact catggcgcca caaggagaac gtggagcaga
 1801 gctgcgaacc tggtcaccag agcccaact gcattcgtt catcgagaac
 1851 ctcacggcgc agcagtaacgg ccactacttc tgcgaggccc aggagggctc
 1901 ctacttccgc gaggctcagc actggcagct gctccccgag gacggcatca
 1951 tggccgagca cctgctgggt catgcctgtc ccctggctgc ctgaattc
 2001 AGCTTGGAGT CGACTCTGCT GAAGAGGAGG AAATTCTCCT TGAAGTTCC
 2051 CTGGTGTCA AAGTAAAGGA GTTGCACCA GACGCACCTC TGTTCACTGG
 2101 TCCGGCGTAT TAAAACACGA TACATTGTTA TTGTTACATT TATTAAGCGC
 2151 TAGATTCTGT GCGTTGTTGA TTTACAGACA ATTGTTGTAC GTATTTAAT
 2201 AATTCACTAA ATTATATAATC TTTAGGGTGG TATGTTAGAG CGAAAATCAA
 2251 ATGATTTCA GCGTCTTAT ATCTGAATT AAATATTTAA TCCTCAATAG
 2301 ATTGTAAAA TAGGTTCGA TTAGTTCAA ACAAGGGTTG TTTTCCGAA
 2351 CCGATGGCTG GACTATCTAA TGGATTTCG CTCAACGCCA CAAAACTTGC
 2401 CAAATCTTGT AGCAGCAATC TAGCTTGTC GATATTGTT TGTGTTTGT
 2451 TTTGTAATAA AGGTTCGACG TCGTTCAAA TATTATGCGC TTTTGTATTT
 2501 CTTTCATCAC TGTCGTAGT GTACAATTGA CTCGACGTAAC ACACGTTAAA
 2551 TAAAGCCTGG ACATATTTAA CATCGGGCGT GTTAGCTTTA TTAGGCGAT
 2601 TATCGTCGTC GTCCCAACCC TCGTCGTTAG AAGTTGCTTC CGAAGACGAT
 2651 TTTGCCATAG CCACACGACG CCTATTAATT GTGTCGGCTA ACACGTCCGC
 2701 GATCAAATTG TAGTTGAGC TTTTGGAAAT TATTCTGAT TGCGGGCGTT
 2751 TTTGGCGGG TTTCAATCTA ACTGTGCCCG ATTTTAATTC AGACAACACG
 2801 TTAGAAAGCG ATGGTGCAGG CGGTGGTAAC ATTCAGACG GCAAATCTAC
 2851 TAATGGCGGC GGTGGTGGAG CTGATGATAA ATCTACCATC GGTGGAGGCG
 2901 CAGGCGGGGC TGGCGGGCGA GGCGGAGGCG GAGGTGGTGG CGGTGATGCA
 2951 GACGGCGGTT TAGGCTAAA TTGTCCTTT CAGGCAACAC AGTCGGCACC

3001 TCAACTATTG TACTGGTTTC GGGCGTATGG TGCAC TCTCA GTACAATCTG
 3051 CTCTGATGCC GCATAGTTAA GCCAGCCCCG ACACCCGCCA ACACCCGCTG
 5 3101 ACGCGCCCTG ACGGGCTTGT CTGCTCCGG CATCCGCTTA CAGACAAGCT
 3151 GTGACCGTCT CCGGGAGCTG CATGTGTCAG AGGTTTCAC CGTCATCACC
 10 3201 GAAACGCGCG AGACGAAAGG GCCTCGTGTACGCCTATT TTATAGGTTA
 3251 ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA
 3301 AATGTGCGCG GAACCCCTAT TTGTTTATT TTCTAAATAC ATTCAAATAT
 15 3351 GTATCCGCTC ATGAGACAAT AACCTGTATA AATGCTTCAA TAATATTGAA
 3401 AAAGGAAGAG TATGAGTATT CAACATTCC GTGTCGCCCT TATTCCCTT
 3451 TTTGCGGCAT TTTGCCTTCC TGTTTTGCT CACCCAGAAA CGCTGGTGAA
 20 3501 AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
 3551 TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTCGCCCG CGAAGAACGT
 3601 TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC
 25 3651 CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
 3701 AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT
 3751 GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA
 30 3801 CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA
 3851 CCGCTTTTT GCACAAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG
 3901 GAACCGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG ACACCACGAT
 35 3951 GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACG GCGAAACTAC
 4001 TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GCGGGATAAA
 4051 GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC
 40 4101 TGATAAATCT GGAGCCGGTG AGCGTGGGTG TCGCGGTATC ATTGCAGCAC
 4151 TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG
 4201 AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC
 4251 CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTAC TCATATATAC
 4301 TTTAGATTGA TTAAAAACTT CATTAAAT TTAAAGGAT CTAGGTGAAG
 4351 ATCCTTTTG ATAATCTCAT GACCAAAATC CCTAACGTG AGTTTCGTT
 50 4401 CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC
 4451 CTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA

4501 CCAGCGGTGG TTTGTTGCC GGATCAAGAG CTACCAACTC TTTTCGAA
 4551 GGTAACTGGC TTCAGCAGAG CGCAGATAAC AAATACTGTT CTTCTAGTGT
 5 4601 AGCCGTAGTT AGGCCACAC TTCAAGAACT CTGTAGCACC GCCTACATAC
 4651 CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
 10 4701 GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC
 4751 GGTGGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG
 15 4801 ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGGCCAC
 4851 GCTTCCCAGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG
 20 4901 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACCG CTGGTATCTT
 4951 TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTGTG
 25 5001 ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCCCT
 5051 TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTCCT
 5101 GCGTTATCCC CTGATTCTGT GGATAACCCT ATTACCGCCT TTGAGTGAGC
 25 5151 TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG
 5201 AGGAAGCATC CTGCACCATC GTCTGCTCAT CCATGACCTG ACCATGCAGA
 5251 GGATGATGCT CGTGACGGTT AACGCCCTGA ATCAGCAACG GCTTGGCGTT
 30 5301 CAGCAGCAGC AGACCATTTC CAATCCGCAC CTCGCGGAAA CCGACATCGC
 5351 AGGCTTCTGC TTCAATCAGC GTGCCGTCGG CGGTGTGCAG TTCAACCACC
 5401 GCACGATAGA GATTGGGAT TTCGGCGCTC CACAGTTCG GGTTTCGAC
 35 5451 GTTCAGACGT AGTGTGACGC GATCGGTATA ACCACCAACGC TCATCGATAA
 5501 TTTCACCGCC GAAAGGCAGC GTGCCGCTGG CGACCTGCGT TTCACCCCTGC
 5551 CATAAAGAAA CTGTTACCCG TAGGTAGTCA CGCAACTCGC CGCACATCTG
 40 5601 AACTTCAGCC TCCAGTACAG CGGGCTGAA ATCATCATTA AAGCGAGTGG
 5651 CAACATGGAA ATCGCTGATT TGTGTAGTCG GTTTATGCAG CAACGAGACG
 45 5701 TCACGGAAAA TGCCGCTCAT CCGCCACATA TCCTGATCTT CCAGATAACT
 5751 GCCGTCACTC CAACGCAGCA CCATCACCGC GAGGCCGGTT TCTCCGGCGC
 5801 GTAAAAATGC GCTCAGGTCA AATTCAACG GCACACGACT GTCCTGGCCG
 5851 TAACCGACCC AGCGCCCGTT GCACCAACAGA TGAAACGCCG AGTTAACGCC
 5901 ATCAAAAATA ATTTCGCGTCT GGCCCTCCTG TAGCCAGCTT TCATCAACAT

5 5951 TAAATGTGAG CGAGTAACAA CCCGTCGGAT TCTCCGTGGG AACAAACGGC
 6001 GGATTGACCG TAATGGGATA GGTCACGTTG GTGTAGATGG GCGCATCGTA
 6051 ACCGTGCATC TGCCAGTTG AGGGGACGAC GACAGTATCG GCCTCAGGAA
 6101 GATCGCACTC CAGCCAGCTT TCCGGCACCG CTTCTGGTGC CGGAAACCAG
 6151 GCAAAGCGCC ATTGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC
 6201 GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCGAAAGGG GGATGTGCTG
 6251 CAAGGCGATT AAGTTGGTA ACGCCAGGGT TTTCCCAGTC ACGACGTTGT
 6301 AAAACGACGG GATCTATCAT TTTAGCAGT GATTCTAATT GCAGCTGCTC
 6351 TTTGATACAA CTAATTTAC GACGACGATG CGAGCTTTA TTCAACCGAG
 6401 CGTGCATGTT TGCAATCGTG CAAGCGTTAT CAATTTTCA TTATCGTATT
 6451 GTTGCACATC AACAGGCTGG ACACCACGTT GAACTCGCCG CAGTTTGCG
 6501 GCAAGTTGGA CCCGCCGCGC ATCCAATGCA AACTTTCCGA CATTCTGTTG
 6551 CCTACGAACG ATTGATTCTT TGTCCATTGA TCGAAGCGAG TGCCTTCGAC
 6601 TTTTCGTGT CCAGTGTGGC TT

30

35

40

45

50

55

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

5

(i) ANMELDER:

- (A) NAME: Hoechst Marion Roussel Deutschland GmbH
- (B) STRASSE: -
- (C) ORT: Frankfurt
- (D) BUNDESLAND: -
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 65926
- (G) TELEPHON: 069-305-7072
- (H) TELEFAX: 069-35-7175
- (I) TELEX: -

10

(ii) ANMELDEBTITEL: Humanes Semaphorin L (H-Sema-L) und
korrespondierende Semaphorine in anderen Spezies

20

(iii) ANZAHL DER SEQUENZEN: 44

25

(iv) COMPUTER-LESBARE FORM:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

(2) INFORMATION ZU SEQ ID NO: 1:

30

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 2636 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

35

(ii) ART DES MOLEKÜLS: DNS (genomisch)

40

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..2636

45

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CGGGGCCACG GGATGACGCC TCCTCCGCC GGACGTGCCG CCCCCAGCGC ACCGCGCGCC	60
CGCGTCCCTG GCCCGCCGGC TCGGTTGGGG CTTCCGCTGC GGCTGCGGCT GCTGCTGCTG	120
CTCTGGCGG CCGCCGCCTC CGCCCAGGGC CACCTAAGGA GCGGACCCCG CATCTTCGCC	180
GTCTGGAAAG GCCATGTAGG GCAGGACCGG GTGGACTTTG GCCAGACTGA GCCGCACACG	240
GTGCTTTCC ACGAGCCAGG CAGCTCCTCT GTGTGGGTGG GAGGACGTGG CAAGGTCTAC	300

55

	CTCTTTGACT TCCCCGAGGG CAAGAACGCA TCTGTGCGCA CGGTGAATAT CGGCTCCACA	360
5	AAGGGGTCT GTCTGGATAA GCGGGACTGC GAGAACTACA TCACTCTCCT GGAGAGGCCGG	420
	AGTGAGGGGC TGCTGGCCTG TGGCACCAAC GCCCCGGCACC CCAGCTGCTG GAACCTGGTG	480
	AATGGCACTG TGGTGCCACT TGGCGAGATG AGAGGCTACG CCCCTTCAG CCCCCGACGAG	540
10	AACTCCCTGG TTCTGTTGA AGGGGACGAG GTGTATTCCA CCATCCGGAA GCAGGAATAC	600
	AATGGGAAGA TCCCTCGGT CCGCCGCATC CGGGGCGAGA GTGAGCTGTA CACCAGTGAT	660
	ACTGTCAATGC AGAACCCACA GTTCATCAA GCCACCATCG TGCACCAAGA CCAGGCTTAC	720
15	GATGACAAGA TCTACTACTT CTTCCGAGAG GACAATCCTG ACAAGAATCC TGAGGCTCCT	780
	CTCAATGTGT CCCGTGTGGC CCAGTTGTGC AGGGGGGAC AGGGTGGGGA AAGTCACTG	840
20	TCAGTCTCCA AGTGGAACAC TTTTCTGAAA GCCATGCTGG TATGCACTG TGCTGCCACC	900
	AACAAGAACT TCAACAGGCT GCAAGACGTC TTCTGCTCC CTGACCCAG CGGCCAGTGG	960
	AGGGACACCA GGGTCTATGG TGTTTCTCC AACCCCTGGA ACTACTCAGC CGTCTGTGTG	1020
25	TATTCCTCG GTGACATTGA CAAGGTCTTC CGTACCTCCT CACTCAAGGG CTACCACTCA	1080
	AGCCTTCCCA ACCCGGGGCC TGGCAAGTGC CTCCCAGACC AGCAGCCGAT ACCCACAGAG	1140
	ACCTTCCAGG TGGCTGACCG TCACCCAGAG GTGGCGCAGA GGGTGGAGCC CATGGGGCCT	1200
30	CTGAAGACGC CATTGTTCCA CTCTAAATAC CACTACCAGA AAGTGGCCGT TCACCCGATG	1260
	CAAGCCAGCC ACGGGGAGAC CTTTCATGTG CTTTACCTAA CTACAGACAG GGGCACTATC	1320
	CACAAGGTGG TGGAACCGGG GGAGCAGGAG CACAGCTTCG CCTTCACAT CATGGAGATC	1380
35	CAGCCCTTCC GCGCGCGGC TGCCATCCAG ACCATGTCGC TGGATGCTGA GCGGAGGAAG	1440
	CTGTATGTGA GCTCCCAGTG GGAGGTGAGC CAGGTGCCCG TGGACCTGTG TGAGGTCTAT	1500
	GGCGGGGGCT GCCACCGTTG CCTCATGTCC CGAGACCCCT ACTGCGGCTG GGACCAGGGC	1560
40	CGCTGCATCT CCATCTACAG CTCCGAACGG TCAGTGTGC AATCCATTAA TCCAGCCGAG	1620
	CCACACAAAGG AGTGTCCCAA CCCCCAAACCA GACAAGGCC CACTGCAGAA GTTTCCCTG	1680
	GCCCCAAACT CTGGCTACTA CCTGAGCTGC CCCATGGAAT CCCGCCACGC CACCTACTCA	1740
45	TGGGCCACA AGGAGAACGT GGAGCAGAGC TGCGAACCTG GTCACCAAGAG CCCCAACTGC	1800
	ATCCCTGTTCA TCGAGAACCT CACGGCGCAG CAGTACGGCC ACTACTTCG CGAGGCCAG	1860
50	GAGGGCTCCT ACTTCCGCGA GGCTCAGCAC TGGCAGCTGC TGCCCGAGGA CGGCATCATG	1920
	GCCGAGCACC TGCTGGGTCA TGCTGTGCC CTGGCTGCCT CCCTCTGGCT GGGGGTGTG	1980

	CCCCACACTCA CTCTTGGCTT GCTGGTCCAC TAGGGCCTCC CGAGGCTGGG CATGCCTCAG	2040
5	GCTTCTGCAG CCCAGGGCAC TAGAACGTCT CACACTCAGA GCCGGCTGGC CCGGGAGCTC	2100
	CTTGCCTGCC ACTTCTTCCA GGGGACAGAA TAACCCAGTG GAGGATGCCA GGCCTGGAGA	2160
	CGTCCAGCCG CAGGCGGCTG CTGGGCCCCA GGTGGCCAC GGATGGTGAG GGGCTGAGAA	2220
10	TGAGGGCACC GACTGTGAAG CTGGGGCATC GATGACCCAA GACTTTATCT TCTGGAAAAT	2280
	ATTTTCAGA CTCCTCAAAC TTGACTAAAT GCAGCGATGC TCCCAGCCCA AGAGCCCATG	2340
	GGTCGGGGAG TGGGTTTGGAG TAGGAGAGCT GGGACTCCAT CTCGACCCCTG GGGCTGAGGC	2400
15	CTGAGTCCTT CTGGACTCTT GGTACCCACA TTGCCTCCTT CCCCTCCCTC TCTCATGGCT	2460
	GGGTGGCTGG TGTTCTGAA GACCCAGGGC TACCCCTGTG CCAGCCCTGT CCTCTGCAGC	2520
20	TCCCTCTCTG GTCCTGGTC CCACAGGACA GCCGCCTTGC ATGTTTATTG AAGGATGTTT	2580
	GCTTTCCGGA CGGAAGGACG GAAAAAGCTC TGAAAAAAAAA AAAAAAAAAA AAAAAAA	2636

(2) INFORMATION ZU SEQ ID NO: 2:

25 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 1195 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

30 (ii) ART DES MOLEKÜLS: DNS (genomisch)

35 (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..1195

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

	CGGGGCTGCG GGATGACGCC TCCTCCTCCC GGACGTGCCG CCCCCAGCGC ACCGCGGCC	60
	CGCGTCCTCA GCCTGCCGGC TCGGTTGGG CTCCCGCTGC GGCTGCGGCT TCTGCTGGTG	120
45	TTCTGGGTGG CCGCCGCCTC CGCCAAGGC CACTCGAGGA GCGGACCCCG CATCTCCGCC	180
	GTCTGGAAAG GGCAGGACCA TGTGGACTTT AGCCAGCCTG AGCCACACAC CGTGCTTTTC	240
	CATGAGCCGG GCAGCTCTC TGTCTGGTG GGTGGACGTG GCAAGGTCTA CCACTTCAAC	300
50	TTCCCCGAGG GCAAGAATGC CTCTGTGCC ACAGGTGAACA TCGGCTCCAC AAAGGGTCC	360
	TGTCAGGACA AACAGGACTG TGGGAATTAC ATCACTCTTC TAGAAAGGCG GGGTAATGGG	420

5	CTGCTGGTCT GTGGCACCAA TGCCCGGAAG CCCAGCTGCT GGAACTTGGT GAATGACAGT	480
	GTGGTGTGATGT CACTTGGTGA GATGAAAGGC TATGCCCT TCAGCCCGA TGAGAACTCC	540
	CTGGTTCTGT TTGAAGGAGA TGAAGTGTAC TCTACCATCC GGAAGCAGGA ATACAACGGG	600
10	AAGATCCCTC GGTTTCGACG CATTGGGGC GAGAGTGAAC TGTACACAAG TGATACAGTC	660
	ATGCAGAACC CACAGTTCAT CAAGGCCACC ATTGTGCACC AAGACCAAGC CTATGATGAT	720
	AAGATCTACT ACTTCTTCCG AGAAGACAAC CCTGACAAGA ACCCCGAGGC TCCTCTCAAT	780
15	GTGTCCCGAG TAGCCCAGTT GTGCAGGGGG GACCAGGGTG GTGAGAGTTC GTTGTCTGTC	840
	TCCAAGTGGA ACACCTTCCT GAAAGCCATG TTGGTCTGCA GCGATGCAGC CACCAACAGG	900
	AACTTCAATC GGCTGCAAGA TGTCTTCCTG CTCCCTGACC CCAGTGGCCA GTGGAGAGAT	960
20	ACCAGGGTCT ATGGCGTTTT CTCCAACCCC TGAAACTACT CAGCTGTCTG CGTGTATTG	1020
	CTTGGTGACA TTGACAGAGT CTTCCGTACC TCATCGCTCA AAGGCTACCA CATGGGCCTT	1080
	TCCAACCCCTC GACCTGGCAT GTGCCTCCCA AAAAAGCAGC CCATACCCAC AGAAACCTTC	1140
25	CAGGTAGCTG ATAGTCACCC AGAGGTGGCT CAGAGGGTGG AACCTATGGG GCCCC	1195

(2) INFORMATION ZU SEQ ID NO: 3:

30	(i) SEQUENZ CHARAKTERISTIKA:
	(A) LÄNGE: 666 Aminosäuren
	(B) ART: Aminosäure
	(C) STRANGFORM: Einzel
	(D) TOPOLOGIE: linear
35	(ii) ART DES MOLEKÜLS: Protein
	(ix) MERKMALE:
40	(A) NAME/SCHLÜSSEL: Protein
	(B) LAGE: 1..666
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:
45	Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala 1 5 10 15
	Arg Val Pro Gly Pro Pro Ala Arg Leu Gly Leu Pro Leu Arg Leu Arg 20 25 30
50	Leu Leu Leu Leu Trp Ala Ala Ala Ser Ala Gln Gly His Leu 35 40 45

EP 0 892 047 A2

Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys Gly His Val Gly Gln
50 55 60

5 Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu Phe His
65 70 75 80

Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr
85 90 95

10 Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn
100 105 110

15 Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu Asn
115 120 125

Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys Gly
130 135 140

20 Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Gly Thr Val
145 150 155 160

Val Pro Leu Gly Glu Met Arg Gly Tyr Ala Pro Phe Ser Pro Asp Glu
165 170 175

25 Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg
180 185 190

30 Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly
195 200 205

Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe
210 215 220

35 Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile
225 230 235 240

Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro
245 250 255

40 Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly
260 265 270

Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met
275 280 285

45 Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn Phe Asn Arg Leu Gln
290 295 300

50 Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg
305 310 315 320

Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val
325 330 335

EP 0 892 047 A2

Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys
 340 345 350
 5 Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro
 355 360 365
 Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His
 370 375 380
 10 Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro
 385 390 395 400
 Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met
 15 405 410 415
 Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp
 420 425 430
 20 Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser
 435 440 445
 Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala
 450 455 460
 25 Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser
 465 470 475 480
 Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr
 485 490 495
 30 Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly
 500 505 510
 Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val
 35 515 520 525
 Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro
 530 535 540
 40 Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser
 545 550 555 560
 Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser
 565 570 575
 45 Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln
 580 585 590
 Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr
 595 600 605
 50 Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala
 610 615 620

55

EP 0 892 047 A2

Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu
625 630 635 640

5 Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu
645 650 655

10 Pro Thr Leu Thr Leu Gly Leu Leu Val His
660 665

(2) INFORMATION ZU SEQ ID NO: 4:

15 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 394 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

20 (ii) ART DES MOLEKÜLS: Protein

25 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: Protein
- (B) LAGE: 1..394

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

30 Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala
1 5 10 15

Arg Val Leu Ser Leu Pro Ala Arg Phe Gly Leu Pro Leu Arg Leu Arg
20 25 30

35 Leu Leu Leu Val Phe Trp Val Ala Ala Ala Ser Ala Gln Gly His Ser
35 40 45

Arg Ser Gly Pro Arg Ile Ser Ala Val Trp Lys Gly Gln Asp His Val
50 55 60

40 Asp Phe Ser Gln Pro Glu Pro His Thr Val Leu Phe His Glu Pro Gly
65 70 75 80

Ser Phe Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr His Phe Asn
85 90 95

45 Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn Ile Gly Ser
100 105 110

Thr Lys Gly Ser Cys Gln Asp Lys Gln Asp Cys Gly Asn Tyr Ile Thr
115 120 125

50 Leu Leu Glu Arg Arg Gly Asn Gly Leu Leu Val Cys Gly Thr Asn Ala
130 135 140

EP 0 892 047 A2

Arg Lys Pro Ser Cys Trp Asn Leu Val Asn Asp Ser Val Val Met Ser
 145 150 155 160
 5 Leu Gly Glu Met Lys Gly Tyr Ala Pro Phe Ser Pro Asp Glu Asn Ser
 165 170 175
 Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg Lys Gln
 180 185 190
 10 Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly Glu Ser
 195 200 205
 Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe Ile Lys
 210 215 220
 Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile Tyr Tyr
 225 230 235 240
 20 Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro Leu Asn
 245 250 255
 Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Glu Ser
 260 265 270
 25 Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met Leu Val
 275 280 285
 Cys Ser Asp Ala Ala Thr Asn Arg Asn Phe Asn Arg Leu Gln Asp Val
 290 295 300
 30 Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg Val Tyr
 305 310 315 320
 35 Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val Tyr Ser
 325 330 335
 Leu Gly Asp Ile Asp Arg Val Phe Arg Thr Ser Ser Leu Lys Gly Tyr
 40 340 345 350
 His Met Gly Leu Ser Asn Pro Arg Pro Gly Met Cys Leu Pro Lys Lys
 355 360 365
 45 Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Ser His Pro Glu
 370 375 380
 Val Ala Gln Arg Val Glu Pro Met Gly Pro
 385 390
 50

55

(2) INFORMATION ZU SEQ ID NO: 5:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 23 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

5

(ii) ART DES MOLEKÜLS: DNS (genomisch)

10

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..23

15

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

ACTCACTATA GGGCTCGAGC GGC

23

20

(2) INFORMATION ZU SEQ ID NO: 6:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

25

(ii) ART DES MOLEKÜLS: DNS (genomisch)

30

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

35

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

AGCCGCACAC GGTGCTTTTC

20

40

(2) INFORMATION ZU SEQ ID NO: 7:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

45

(ii) ART DES MOLEKÜLS: DNS (genomisch)

50

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

GCACAGATGC GTTCTTGCCC

5

20

(2) INFORMATION ZU SEQ ID NO: 8:

(i) SEQUENZ CHARAKTERISTIKA:

10 (A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

15 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

20 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

ACCATAGACC CTGGTGTCCC

25

20

(2) INFORMATION ZU SEQ ID NO: 9:

(i) SEQUENZ CHARAKTERISTIKA:

30 (A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

35 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

40 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

GCAGTGATGC TGCCACCAAC

45

20

(2) INFORMATION ZU SEQ ID NO: 10:

(i) SEQUENZ CHARAKTERISTIKA:

50 (A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure

55

(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon
10 (B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

15 CCAGACCATG TCGCTGGATG

20

(2) INFORMATION ZU SEQ ID NO: 11:

20 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

30 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

35 ACATGAGGCA ACCGTGGCAG

20

(2) INFORMATION ZU SEQ ID NO: 12:

40 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 27 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

50 CCATCCTAAT ACGACTCACT ATAGGGC

27

55

(2) INFORMATION ZU SEQ ID NO: 13:

5 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: DNS (genomisch)

15 (ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20

20 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

AGGTAGACCT TGCCACGTCC

20

(2) INFORMATION ZU SEQ ID NO: 14:

25 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 23 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

30 (ii) ART DES MOLEKÜLS: DNS (genomisch)

35 (ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..23

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

GAACTTCAAC AGGCTGCAAG ACG

23

(2) INFORMATION ZU SEQ ID NO: 15:

45 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

50 (ii) ART DES MOLEKÜLS: DNS (genomisch)

55

5
(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20

10
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

ATGCTGAGCG GAGGAAGCTG

20

15
(2) INFORMATION ZU SEQ ID NO: 16:

15
(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

20
(ii) ART DES MOLEKÜLS: DNS (genomisch)

25
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

CCGCCATACA CCTCACACAG

20

25
(2) INFORMATION ZU SEQ ID NO: 17:

30
(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 28 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

35
(ii) ART DES MOLEKÜLS: DNS (genomisch)

35
(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..28

40

40
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

45
CTGGAAGCTT TCTGTGGGTA TCGGCTGC

28

45
(2) INFORMATION ZU SEQ ID NO: 18:

50
(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 25 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

55

5 (ii) ART DES MOLEKÜLS: DNS (genomisch)

10 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..25

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

20 TTTGGATCCC TGGTTCTGTT TGAAG

25

15 (2) INFORMATION ZU SEQ ID NO: 19:

20 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 50 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: DNS (genomisch)

30 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..50

35 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

40 TTCTAGAATT CAGCGGCCGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

50

35 (2) INFORMATION ZU SEQ ID NO: 20:

40 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 27 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: DNS (genomisch)

50 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..27

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

60 GGGGAAAGTT CACTGTCAGT CTCCAAG

27

55

(2) INFORMATION ZU SEQ ID NO: 21:

5 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 26 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

15 GGGAAATACAC ACAGACGGCT GAGTAG

26

(2) INFORMATION ZU SEQ ID NO: 22:

20 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 22 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: DNS (genomisch)

30 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..22

35 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

35 AGCAAGTTCA GCCTGGTTAA GT

22

(2) INFORMATION ZU SEQ ID NO: 23:

40 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 21 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: DNS (genomisch)

50 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..21

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

TTATGAGTAT TTCTTCCAGG G

21

5

(2) INFORMATION ZU SEQ ID NO: 24:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 26 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

15

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..26

20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

CCATTAATCC AGCCGAGCCA CACAAG

26

25

(2) INFORMATION ZU SEQ ID NO: 25:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 25 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

35

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..25

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

CATCTACAGC TCCGAACGGT CAGTG

25

45

(2) INFORMATION ZU SEQ ID NO: 26:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

55

EP 0 892 047 A2

(ii) ART DES MOLEKÜLS: DNS (genomisch)

5 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

CAGCGGAAGC CCCAACCGAG

20

15 (2) INFORMATION ZU SEQ ID NO: 27:

20 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 23 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: DNS (genomisch)

30 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..23

35 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

30 GGGATGACGC CTCCTCCGCC CGG

23

40 (2) INFORMATION ZU SEQ ID NO: 28:

35 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 31 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: DNS (genomisch)

45 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..31

50 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

50 AAGCTTCACG TGGACCAGCA AGCCAAGAGT G

31

55

(2) INFORMATION ZU SEQ ID NO: 29:

5 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 25 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: DNS (genomisch)

15 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

20 AAGCTTTTTC CGTCCTTCCG TCCGG

25

(2) INFORMATION ZU SEQ ID NO: 30:

25 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 24 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

30 (ii) ART DES MOLEKÜLS: DNS (genomisch)

35 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..24

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

40 ATGGTGAGCA AGGGCGAGGA GCTG

24

(2) INFORMATION ZU SEQ ID NO: 31:

45

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 24 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

50

(ii) ART DES MOLEKÜLS: DNS (genomisch)

55

5
(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..24

5

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

10 CTTGTACAGC TCGTCCATGC CGAG

24

10

(2) INFORMATION ZU SEQ ID NO: 32:

15

(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 25 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

20

(ii) ART DES MOLEKÜLS: DNS (genomisch)

25

(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..25

30

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

25

30

(2) INFORMATION ZU SEQ ID NO: 33:

35

(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 25 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

40

(ii) ART DES MOLEKÜLS: DNS (genomisch)

45

(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..25

50

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

25

50

GAGCGATGAG GTACGGAAGA CTCTG

55

(2) INFORMATION ZU SEQ ID NO: 34:

(i) SEQUENZ CHARAKTERISTIKA:

5 (A) LÄNGE: 5856 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

15 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..5856

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

20	AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC	60
	ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	120
	TCACTCATTA GGCACCCAG GCTTTACACT TTATGTTCC GGCTCGTATG TTGTGTGGAA	180
25	TTGTGAGCGG ATAACAATT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTTC	240
	ACGTGGACCA GCAAGCCAAG AGTGAGTGTG GGCAGCACCC CCAGCCAGAG GGAGGCAGCC	300
	AGGGCACAGG CATGACCCAG CAGGTGCTCG GCCATGATGC CGTCCTCGGG CAGCAGCTGC	360
30	CAGTGCTGAG CCTCGGGAA GTAGGAGCCC TCCTGGCCT CGCAGAAAGTA GTGGCCGTAC	420
	TGCTGCGCCG TGAGGTTCTC GATGAACAGG ATGCAGTTGG GGCTCTGGTG ACCAGGTTCG	480
	CAGCTCTGCT CCACGTTCTC CTTGTCGGC CATGAGTAGG TGGCGTGGCG GGATTCCATG	540
35	GGGCAGCTCA GGTAGTAGCG AGAGTTGGG GCCAGGGAAA CCTTCTGCAG TGGGGCTTG	600
	TCTGGTTTGG GGTTGGGACA CTCCCTGTGT GGCTCGGCTG GATTAATGGA TTGCAGCACT	660
	GACCGTTCGG AGCTGTAGAT GGAGATGCAG CGGCCCTGGT CCCAGCCGCA GTAGGGGTCT	720
40	CGGGACATGA GGCAACCGTG GCAGCCCCCG CCATAGACCT CACACAGGTC CAGGGGCACC	780
	TGGCTCACCT CCCACTGGGA GCTCACATAC AGCTTCTCC GCTCAGCATC CAGCGACATG	840
45	GTCTGGATGG CAGCCGCGCG CGGGAAAGGGC TGGATCTCCA TGATGTTGAA GGCGAAGCTG	900
	TGCTCCTGCT CCCCCGGTTC CACCACTTG TGGATAGTGC CCCTGTCTGT AGTTAGGTAA	960
	AGCACATGAA AGGTCTCCCC GTGGCTGGCT TGCATGCGGT GAACGGCCAC TTTCTGGTAG	1020
50	TGGTATTTAG AGTGGAACAA TGGCGTCTTC AGAGGCCCCA TGGGCTCCAC CCTCTGCGCC	1080
	ACCTCTGGGT GACGGTCAGC CACCTGGAAG GTCTCTGTGG GTATCGGCTG CTGGTCTGGG	1140

	AGGCACTTGC CAGGCCGCGG GTTGGGAAGG CTTGAGTGGT AGCCCTTGAG TGAGGAGGTA	1200
5	CGGAAGACCT TGTCAATGTC ACCGAGGGAA TACACACAGA CGGCTGAGTA GTTCCAGGG	1260
	TTGGAGAAAA CACCATAGAC CCTGGTGTCC CTCCACTGGC CGCTGGGTC AGGGAGCAGG	1320
	AAGACGTCTT GCAGCCTGTT GAAGTTCTTG TTGGTGGCAG CATCACTGCA TACCAGCATG	1380
10	GCTTCAGAA AAGTGTCCA CTTGGAGACT GACAGTGAAC TTTCCCCACC CTGGTCCCC	1440
	CTGCACAACT GGGCCACACG GGACACATTG AGAGGAGCCT CAGGATTCTT GTCAGGATTG	1500
	TCCTCTCGGA AGAAGTAGTA GATCTTGTCA TCGTAAGCCT GGTCTTGGTG CACGATGGTG	1560
15	GCTTGATGA ACTGTGGTT CTGCATGACA GTATCACTGG TGTACAGCTC ACTCTCGCCC	1620
	CGGATGCGGC GGAACCGAGG GATCTCCCA TTGTATTCCCT GCTTCCGGAT GGTGGAATAC	1680
20	ACCTCGTCCC CTTCAAACAG AACCAAGGGAG TTCTCGTCCG GGCTGAAGGG GGCGTAGCCT	1740
	CTCATCTCGC CAAGTGGCAC CACAGTGCCTA TTCAACAGGT TCCAGCAGCT GGGGTGCCGG	1800
	GCGTTGGTGC CACAGGCCAG CAGCCCTCA CTCCGCCTCT CCAGGAGAGT GATGTAGTTC	1860
25	TCGCAGTCCC GCTTATCCAG ACAGGACCCC TTTGTGGAGC CGATATTCAAC CGTGCACACA	1920
	GATGCGTTCT TGCCCTCGGG GAAGTCAAAG AGGTAGACCT TGCCACGTCC TCCCACCCAC	1980
	ACAGAGGAGC TGCCCTGGCTC GTGGAAAAGC ACCGTGTGCG GTCAGTCTG GCCAAAGTCC	2040
30	ACCCGGTCCT GCCCTACATG GCCTTCCAG ACGGCGAAGA TGCGGGTCC GCTCCTTAGG	2100
	TGGCCCTGGG CGGAGGCGGC GGCCGCCAG AGCAGCAGCA GCAGCCGCAG CCGCAGCGGA	2160
35	AGCCCCAACCC GAGCCGGCGG GCCAGGGACG CGGGCGCGCG GTGCGCTGGG GGCGGCACGT	2220
	CCGGGCGGAG GAGGCGTCAT CCCAACCGA ATTCTGCAGA TATCCATCAC ACTGGCGGCC	2280
	GCTCGAGCAT GCATCTAGAG GGCCCAATTG GCCCTATAGT GAGTCGTATT ACAATTCACT	2340
40	GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCCTGGC GTTACCCAAC TTAATCGCCT	2400
	TGGCAGCACAT CCCCCCTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	2460
	TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG	2520
45	CGCCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC	2580
	CGCTCCCTTC GCTTTCTTCC CTTCCCTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC	2640
50	TCTAAATCGG GGGCTCCCTT TAGGGTCCG ATTTAGAGCT TTACGGCACC TCGACCGCAA	2700
	AAAACTTGAT TTGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG	2760

	CCCTTTGACG TTGGAGTCCA CGTTCTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC	2820
5	ACTCAACCT ATCGCGTCT ATTCTTTGA TTTATAAGGG ATTTGCCGA TTTCGGCCTA	2880
	TTGGTTAAAA AATGAGCTGA TTTAACAAAT TCAGGGCGCA AGGGCTGCTA AAGGAACCGG	2940
	AACACGTAGA AAGCCAGTCC GCAGAAACGG TGCTGACCCC GGATGAATGT CAGCTACTGG	3000
10	GCTATCTGGA CAAGGGAAAA CGCAAGCGCA AAGAGAAAGC AGGTAGCTTG CAGTGGGCTT	3060
	ACATGGCGAT AGCTAGACTG GGCGGTTTA TGGACAGCAA GCGAACCGGA ATTGCCAGCT	3120
	GGGGCGCCCT CTGGTAAGGT TGGGAAGCCC TGCAAAGTAA ACTGGATGGC TTTCTGCCG	3180
15	CCAAGGATCT GATGGCGCAG GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTT	3240
	CGCATGATTG AACAAAGATGG ATTGCACGCA GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA	3300
	TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT GTCGGCTG	3360
20	TCAGCGCAGG GGCGCCCGT TCTTTTGTC AAGACCGACC TGTCCGGTGC CCTGAATGAA	3420
	CTGCAGGACG AGGCAGCGCG GCTATCGTGG CTGGCACGA CGGGCGTTCC TTGCGCAGCT	3480
	GTGCTCGACG TTGTCACTGA AGCGGGAAGG GACTGGCTGC TATTGGCGA AGTCCGGGG	3540
25	CAGGATCTCC TGTATCTCG CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA	3600
	ATGCGGCGGC TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA AGCGAACAT	3660
	CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG TCGATCAGGA TGATCTGGAC	3720
30	GAAGAGCATC AGGGGCTCGC GCCAGCCGAA CTGTCGCCA GGCTCAAGGC GCGCATGCC	3780
	GACGGCGAGG ATCTCGTCGT GATCCATGGC GATGCCCTGCT TGCCGAATAT CATGGTGGAA	3840
35	AATGGCCGCT TTTCTGGATT CAACGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG	3900
	GACATAGCGT TGGATACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG GGCTGACCGC	3960
	TTCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTCGCAGC GCATGCCCTT CTATGCCCTT	4020
40	CTTGACGAGT TCTTCTGAAT TGAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG	4080
	CCCTTATTCC CTTTTTGCG GCATTTGCC TTCCCTTTTG TGCTCACCA GAAACGCTGG	4140
	TGAAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC	4200
45	TCAACAGCGG TAAGATCCTT GAGAGTTTC GCCCCGAAGA ACGTTTCCA ATGATGAGCA	4260
	CTTTTAAAGT TCTGCTATGT CATAACTAT TATCCGTAT TGACGCCGGG CAAGAGCAAC	4320
50	TCGGTCGCCG GGCGCGGTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA	4380
	AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG	4440

5	ATAACACTGC GGCCAACCTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT	4500
	TTTGCACAA CATGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG	4560
10	AAGCCATACC AAACGACGAG AGTGACACCA CGATGCCTGT AGCAATGCCA ACAACGTTGC	4620
	GCAAACATTAA AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTAA ATAGACTGGA	4680
15	TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCGCTCGGC CCTTCCGGCT GGCTGGTTTA	4740
	TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC	4800
20	CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG	4860
	ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT	4920
25	CAGACCAAGT TTACTCATAT ATACTTTAGA TTGATTAAA ACTTCATTAA TAATTTAAAA	4980
	GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTT	5040
30	CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT	5100
	TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCAACC GCTACCAGCG GTGGTTTGT	5160
35	TGCCGGATCA AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA	5220
	TACCAAATAC TGTCCCTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG	5280
40	CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC AGTGGCGATA	5340
	AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTGCG	5400
45	GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACGTGA	5460
	GATACTACAC GCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA	5520
	GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA	5580
	ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTGCCCACCT CTGACTTGAG CGTCGATTTT	5640
50	TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC	5700
	GGTTCCCTGGC CTTTTGCTGG CCTTTGCTC ACATGTTCTT TCCCTGCGTTA TCCCCTGATT	5760
	CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA	5820
	CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG CGGAAG	5856

(2) INFORMATION ZU SEQ ID NO: 35:

55 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 7475 Basenpaare
 (B) ART: Nukleinsäure

(C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 10 (B) LAGE: 1..7475

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35:

15	GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC TGCTCTGATG	60
	CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG	120
	CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC	180
20	TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT	240
	GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTTCAT AGCCCATATA	300
	TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC	360
25	CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCAATA GGAACTTTCC	420
	ATTGACGTCA ATGGGTGGAC TATTACGGT AAACGTGCCA CTTGGCAGTA CATCAAGTGT	480
	ATCATATGCC AAGTACGCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT	540
30	ATGCCAGTA CATGACCTTA TGGGACTTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA	600
	TCGCTATTAC CATGGTGTG CGGTTTTGGC AGTACATCAA TGGCGTGGA TAGCGGTTG	660
35	ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTG TTTTGGCACC	720
	AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGCG	780
	GTAGGCGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCA	840
40	CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCAA GCTGGCTAGC	900
	GTTTAAACGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT GCAGAATTG	960
45	GCTTGGGATG ACGCCTCTC CGCCCGGACG TGCCGCCCCC AGCGCACCGC GCGCCCGCGT	1020
	CCCTGGCCCG CGGGCTCGGT TGGGGCTTCC GCTGGGGCTG CGGCTGCTGC TGCTGCTCTG	1080
	GGCGGCCGCC GCCTCCGCC AGGGCCACCT AAGGAGCGGA CCCCGCATCT TCGCCGTCTG	1140
50	GAAAGGCCAT GTAGGGCAGG ACCGGGTGGA CTTTGGCCAG ACTGAGCCGC ACACGGTGCT	1200
	TTTCCACGAG CCAGGCAGCT CCTCTGTGTG GGTGGGAGGA CGTGGCAAGG TCTACCTCTT	1260

55

EP 0 892 047 A2

	TGACTTCCCC GAGGGCAAGA ACGCATCTGT GCGCACGGTG AATATCGGCT CCACAAAGGG	1320
5	GTCTCTGCTG GATAAGCGGG ACTGCGAGAA CTACATCACT CTCTCTGGAGA GGCGGAGTGA	1380
	GGGGCTGCTG GCCTGTGGCA CCAACGCCCC GCACCCCCAGC TGCTGGAACC TGGTGAATGG	1440
	CACTGTGGTG CCACTTGGCG AGATGAGAGG CTACGCCCCC TTCAGCCCCG ACGAGAACTC	1500
10	CCTGGTTCTG TTTGAAGGGG ACGAGGTGTA TTCCACCAC CGGAAGCAGG AATACAATGG	1560
	GAAGATCCCT CGGTTCCGCC GCATCCGGGG CGAGAGTGAG CTGTACACCA GTGATACTGT	1620
	CATGCAGAAC CCACAGTTCA TCAAAGCCAC CATCGTGCAC CAAGACCAGG CTTACGATGA	1680
15	CAAGATCTAC TACTTCTTCC GAGAGGACAA TCCTGACAAG AATCCTGAGG CTCCTCTCAA	1740
	TGTGTCCCGT GTGGCCCAAGT TGTGCAGGGG GGACCAAGGT GGGGAAAGTT CACTGTCAGT	1800
	CTCCAAGTGG AACACTTTTC TGAAAGCCAT GCTGGTATGC AGTGATGCTG CCACCAACAA	1860
20	GAACATTCAAC AGGCTGCAAG ACGTCTCCT GCTCCCTGAC CCCAGCGGCC AGTGGAGGG	1920
	CACCAAGGTC TATGGTGTTC TCTCCAACCC CTGGAACCTAC TCAGCCGTCT GTGTGTATTTC	1980
	CCTCGGTGAC ATTGACAAGG TCTTCCGTAC CTCCTCACTC AAGGGCTACC ACTCAAGCCT	2040
25	TCCCAACCCG CGGCCTGGCA AGTGCCTCCC AGACCAAGCAG CGGATACCCA CAGAGACCTT	2100
	CCAGGTGGCT GACCGTCACC CAGAGGTGGC GCAGAGGGTG GAGCCCATGG GGCCCTCTGAA	2160
	GACGCCATTG TTCCACTCTA AATACCACTA CCAGAAAGTG GCCGTTCAACC GCATGCAAGC	2220
30	CAGCCACGGG GAGACCTTTC ATGTGCTTTA CCTAACTACA GACAGGGGCA CTATCCACAA	2280
	GGTGGTGGAA CGGGGGGAGC AGGAGCACAG CTTGCCCTTC AACATCATGG AGATCCAGCC	2340
35	CTTCCGCCGC GCGGCTGCCA TCCAGACCAT GTCGCTGGAT GCTGAGCGGA GGAAGCTGTA	2400
	TGTGAGCTCC CAGTGGGAGG TGAGCCAGGT GCCCCTGGAC CTGTGTGAGG TCTATGGCGG	2460
	GGGCTGCCAC GGTTGCCTCA TGTCCCCAGA CCCCTACTGC GGCTGGGACC AGGGCCGCTG	2520
40	CATCTCCATC TACAGCTCCG AACGGTCAGT GCTGCAATCC ATTAATCCAG CCGAGCCACA	2580
	CAAGGAGTGT CCCAACCCCA AACCAGACAA GGCCCCACTG CAGAAGGTTT CCCCTGGCCCC	2640
	AAACTCTCGC TACTACCTGA GCTGCCCAT GGAATCCCAG CACGCCACCT ACTCATGGCG	2700
45	CCACAAGGAG AACGTGGAGC AGAGCTGCGA ACCTGGTCAC CAGAGCCCCA ACTGCATCCT	2760
	GTTCATCGAG AACCTCACCG CGCAGCAGTA CGGCCACTAC TTCTGCGAGG CCCAGGAGGG	2820
50	CTCCTACTTC CGCGAGGCTC AGCACTGGCA GCTGCTGCCG GAGGACGGCA TCATGGCCGA	2880
	GCACCTGCTG GGTCACTGCCT GTGCCCTGGC TGCCCTCCCTC TGGCTGGGGG TGCTGCCAC	2940

	ACTCACTCTT GGCTTGCTGG TCCACGTGAA GCTTGGGCCA GAACAAAAAC TCATCTCAGA	3000
5	AGAGGATCTG AATAGCGCCG TCGACCATCA TCATCATCAT CATTGAGTTT AAACCGCTGA	3060
	TCAGCCTCGA CTGTGCCCTTC TAGTTGCCAG CCATCTGTTG TTTGCCCTC CCCCCTGCCCT	3120
	TCCTTGACCC TGGAAGGTGC CACTCCCCT GTCCTTCTCCT AATAAAATGA GGAAATTGCA	3180
10	TCGCATTGTC TGAGTAGGTG TCATTCTATT CTGGGGGTG GGGTGGGCA GGACAGCAAG	3240
	GGGGAGGATT GGGAAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGCTC TATGGCTTCT	3300
	GAGGCGGAAA GAACCAGCTG GGGCTCTAGG GGGTATCCCC ACGCGCCCTG TAGCGGCGCA	3360
15	TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA	3420
	GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT	3480
	CAAGCTCTAA ATCGGGGCAT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC	3540
20	CCCAAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT	3600
	TTTCGCCCTT TGACGTTGGA GTCCACGTTG TTTAATAGTG GACTCTTGTG CCAAACGTGGA	3660
25	ACAACACTCA ACCCTATCTC GGCTTATTCT TTTGATTAA AAGGGATTTT GGGGATTTCG	3720
	GCCTATTGGT TAAAAAAATGA GCTGATTAA CAAAATTTA ACGCGAATTA ATTCTGTGGA	3780
	ATGTGTGTCA GTTAGGGTGT GGAAAGTCCC CAGGCTCCCC AGGCAGGCAG AAGTATGCAA	3840
30	AGCATGCATC TCAATTAGTC AGCAACCAGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC	3900
	AGAAGTATGCC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC CCTAACTCCG	3960
35	CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC CGCCCCATGG CTGACTAATT	4020
	TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCTGCCTCTG AGCTATTCCA GAAGTAGTGA	4080
	GGAGGCTTTT TTGGAGGCCT AGGCTTTGC AAAAGCTCC CGGGAGCTTG TATATCCATT	4140
40	TTCGGATCTG ATCAAGAGAC AGGATGAGGA TCGTTTCGCA TGATTGAACA AGATGGATTG	4200
	CACCGAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCTG GCTATGACTG GGCACAAACAG	4260
	ACAATCGGCT GCTCTGATGC CGCCGTGTTG CGGCTGTCAAG CGCAGGGCGC CCCGGTTCTT	4320
45	TTTGTCAAGA CCGACCTGTC CGGTGCCCTG AATGAACCTGC AGGACGAGGC AGCGCGGCTA	4380
	TCGTGGCTGG CCACGACGGG CGTTCCCTGC GCAGCTGTGC TCGACGTTGT CACTGAAGCG	4440
50	GGAAAGGGACT GGCTGCTATT GGGCGAAGTG CGGGGGCAGG ATCTCCTGTC ATCTCACCTT	4500
	GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA TACGCTTGAT	4560

	CCGGCTACCT	GCCCATTGCA	CCACCAAGCG	AAACATCGA	TCGAGCGAGC	ACGTACTCGG	4620
5	ATGGAAGCGG	GTCTTGTGCA	TCAGGATGAT	CTGGACGAAG	AGCATCAGGG	GCTCGCGCCA	4680
	GCCGAACTGT	TCGCCAGGCT	CAAGGCGCGC	ATGCCGACG	GCGAGGATCT	CGTCGTGACC	4740
	CATGGCGATG	CCTGCTTGCC	GAATATCATG	GTGGAAAATG	GCCGCTTTTC	TGGATTCATC	4800
10	GACTGTGGCC	GGCTGGGTGT	GGCGGACCGC	TATCAGGACA	TAGCGTTGGC	TACCCGTGAT	4860
	ATTGCTGAAG	AGCTTGGCGG	CGAATGGCT	GACCGCTTCC	TCGTGCTTTA	CGGTATCGCC	4920
	GCTCCCGATT	CGCAGCGCAT	CGCCTCTAT	CGCCTCTTG	ACGAGTTCTT	CTGAGCGGGA	4980
15	CTCTGGGGTT	CGAAATGACC	GACCAAGCGA	CGCCCAACCT	GCCATCACGA	GATTCGATT	5040
	CCACCGCCCG	CTTCTATGAA	AGGTTGGCT	TCGGAATCGT	TTTCCGGGAC	GCCGGCTGGA	5100
	TGATCCTCCA	GCGCGGGGAT	CTCATGCTGG	AGTTCTTCGC	CCACCCCCAAC	TTGTTTATTG	5160
20	CAGCTTATAA	TGGTTACAAA	TAAAGCAATA	GCATCACAAA	TTTCACAAAT	AAAGCATTTC	5220
	TTTCACTGCA	TTCTAGTTGT	GGTTTGTCCA	AACTCATCAA	TGTATCTTAT	CATGTCTGTA	5280
	TACCGTCGAC	CTCTAGCTAG	AGCTTGGCGT	AATCATGGTC	ATAGCTGTT	CCTGTGTGAA	5340
25	ATTGTTATCC	GCTCACAAATT	CCACACAAACA	TACGAGCGGG	AAGCATAAAG	TGTAAAGCCT	5400
	GGGGTGCCTA	ATGAGTGAGC	TAACTCACAT	TAATTGCGTT	GCGCTCACTG	CCCGCTTTCC	5460
	AGTCGGGAAA	CCTGTCGTG	CAGCTGCATT	AATGAATCGG	CCAACCGCGG	GGGAGAGGCG	5520
30	GTGCGTAT	TGGCGCTCT	TCCGCTTCCT	CGCTCACTGA	CTCGCTGCGC	TCGGTGTTC	5580
	GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAT	ACGGTTATCC	ACAGAACATCAG	5640
35	GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	AAGGCCAGCA	AAAGGCCAGG	AACCGTAAAA	5700
	AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC	TCCGCCCCCCC	TGACGAGCAT	CACAAAATC	5760
	GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA	CAGGACTATA	AAGATACCAAG	GCGTTTCCCC	5820
40	CTGGAAGCTC	CCTCGTGC	TCTCCTGTT	CGACCCCTGCC	GCTTACCGGA	TACCTGTCCG	5880
	CCTTTCTCCC	TCGGGAAGC	GTGGCGCTTT	CTCAATGCTC	ACGCTGTAGG	TATCTCAGTT	5940
45	CGGTGTAGGT	CGTTCGCTCC	AAGCTGGCT	GTGTGACGA	ACCCCCCGTT	CAGCCCGACC	6000
	GCTGCGCCTT	ATCCGGTAAC	TATCGTCTTG	AGTCCAACCC	GGTAAGACAC	GACTTATCGC	6060
	CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGTACAG	6120
50	AGTTCTTGAA	GTGGTGGCCT	AACTACGGCT	ACACTAGAAG	GACAGTATT	GGTATCTGCG	6180
	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA	6240

5	CCACCGCTGG TAGCGCTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG	6300
10	GATCTCAAGA AGATCCTTG ATCTTTCTA CGGGTCTGA CGCTCAGTGG AACGAAACT	6360
15	CACGTTAAGG GATTTGGTC ATGAGATTAT CAAAAAGGAT CTPCACCTAG ATCCTTTAA	6420
20	ATTAAAAATG AAGTTTAAA TCAATCTAA GTATATATGA GTAAACTTGG TCTGACAGTT	6480
25	ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTCGT TCATCCATAG	6540
30	TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA	6600
35	G TGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAACC	6660
40	AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT	6720
45	CTATTAATTG TTGCCGGAA GCTAGAGTAA GTAGTTGCC AGTTAATAGT TTGCGCAACG	6780
50	TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA	6840
55	GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG	6900
60	TTAGCTCCTT CGGTCCCTCG ATCGTTGTCA GAAGTAAGTT GGCGCAGTG TTATCACTCA	6960
65	TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTCTG	7020
70	TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGCGA CCGAGTTGCT	7080
75	CTTGGCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA	7140
80	TCATTGGAAA ACGTTCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA	7200
85	GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTACT TTCACCAGCG	7260
90	TTTCTGGGTG AGCAAAAACA GGAAGGCAAATGCGAAA AAAGGGAAATA AGGGCGACAC	7320
95	GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT	7380
100	ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGTTC	7440
105	CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTC	7475

(2) INFORMATION ZU SEQ ID NO: 36:

45 (i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 8192 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

50 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:
 (A) NAME/Schlüssel: exon
 (B) LAGE: 1..8192

5

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36:

10	GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCACTCT CAGTACAATC TGCTCTGATG	60
15	CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG	120
20	CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC	180
25	TTAGGGTTAG GCGTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATAACG CGTTGACATT	240
30	GATTATTGAC TAGTTTAA TAGTAATCAA TTACGGGTC ATTAGTTCAT AGCCCATATA	300
35	TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC	360
40	CCCGCCCAT T GACGTCAATA ATGACGTATG TTCCCATAGT AACGCAATA GGGACTTTCC	420
45	ATTGACGTCA ATGGGTGGAC TATTTACGGT AACTGCCC CTTGGCAGTA CATCAAGTGT	480
50	ATCATATGCC AAGTACGCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT	540
55	ATGCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA	600
60	TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG	660
65	ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTGGCACC	720
70	AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGCG	780
75	GTAGGGCGTGT ACGGTGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA	840
80	CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCAA GCTGGCTAGC	900
85	GTTAACCGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT GCAGAATTG	960
90	GCTTGGGATG ACGCCTCCTC CGCCCGGAGC TGCCGCCCCC AGCGCACCGC GCGCCCGCGT	1020
95	CCCTGGCCCG CGGCTCGGT TGGGCTTCC GCTGCGGCTG CGGCTGCTGC TGCTGCTCTG	1080
100	GGCGGCCGCC GCCTCCGCC AGGGCACCT ARGGAGCGGA CCCCGCATCT TCGCCGTCTG	1140
105	GAAAGGCCAT GTAGGGCAGG ACCGGGTGGA CTTTGGCCAG ACTGAGCCGC ACACGGTGCT	1200
110	TTTCCACGAG CCAGGCAGCT CCTCTGTGTG GGTGGGAGGA CGTGGCAAGG TCTACCTCTT	1260
115	TGACTTCCCC GAGGGCAAGA ACGCATCTGT GCGCACGGTG AATATCGGCT CCACAAAGGG	1320
120	GTCCTGTCTG GATAAGCGGG ACTGCGAGAA CTACATCACT CTCCCTGGAGA GGC GGAGTGA	1380
125	GGGGCTGCTG GCCTGTGGCA CCAACGCCCG GCACCCCAGC TGCTGGAACC TGGTGAATGG	1440

55

EP 0 892 047 A2

	CACTGTGGTG CCACTTGGCG AGATGAGAGG CTACGCCCCC TTCAGCCCGG ACGAGAACTC	1500
5	CCTGGTTCTG TTTGAAGGGG ACGAGGTGA TTCCACCATC CGGAAGCAGG AATACAATGG	1560
	GAAGATCCCT CGGTTCCGCC GCATCCGGGG CGAGAGTGAG CTGTACACCA GTGATACTGT	1620
	CATGCAGAAC CCACAGTTCA TCAAAGCCAC CATCGTCAC CAAGACCAGG CTTACGATGA	1680
10	CAAGATCTAC TACTTCTTCC GAGAGGACAA TCCTGACAAG AATCCTGAGG CTCCCTCAA	1740
	TGTGTCCCGT GTGGCCCACT TGTGCAGGGG GGACCAGGGT GGGGAAAGTT CACTGTCAGT	1800
	CTCCAAGTGG AACACTTTTC TGAAAGCCAT GCTGGTATGC AGTGATGCTG CCACCAACAA	1860
15	GAACATTCAAC AGGCTGCAAG ACGTCTCCT GCTCCCTGAC CCCAGCGGCC AGTGGAGGGA	1920
	CACCAGGGTC TATGGTGTTC TCTCCAACCC CTGGAACTAC TCAGCCGTCT GTGTGTATTIC	1980
	CCTCGGTGAC ATTGACAAGG TCTTCCGTAC CTCCCTCACTC AAGGGCTACC ACTCAAGCCT	2040
20	TCCCAACCCG CGGCCTGGCA AGTGCCTCCC AGACCAGCAG CCGATACCCA CAGAGACCTT	2100
	CCAGGGTGGCT GACCGTCACC CAGAGGTGGC GCAGAGGGTG GAGCCCATGG GGCCTCTGAA	2160
	GACGCCATTG TTCCACTCTA AATACCACTA CCAGAAAGTG GCCGTTCACCC GCATGCAAGC	2220
25	CAGCCACGGG GAGACCTTTC ATGTGCTTTA CCTAACTACA GACAGGGCA CTATCCACAA	2280
	GGTGGTGGAA CGGGGGAGC AGGAGCACAG CTTCGCCTTC AACATCATGG AGATCCAGCC	2340
	CTTCCGCCGC GCGGCTGCCA TCCAGACCAT GTCGCTGGAT GCTGAGGGGA GGAAGCTGTA	2400
30	TGTGAGCTCC CAGTGGGAGG TGAGCCAGGT GCCCCTGGAC CTGTGTGAGG TCTATGGCGG	2460
	GGGCTGCCAC GGTTGCCCTA TGTCGGAGA CCCCTACTGC GGCTGGGACC AGGGCCGCTG	2520
35	CATCTCCATC TACAGCTCG AACGGTCAGT GCTGCAATCC ATTAATCCAG CCGAGCCACA	2580
	CAAGGAGTGT CCCAACCCCA ACCAGACAA GGCCCCACTG CAGAAGGTTT CCCTGGCCCC	2640
	AAACTCTCGC TACTACCTGA GCTGCCCAT GGAATCCCGC CACGCCACCT ACTCATGGCG	2700
40	CCACAAGGAG AACGTGGAGC AGAGCTGCGA ACCTGGTCAC CAGAGCCCCA ACTGCATCCT	2760
	GTTCATCGAG AACCTCACGG CGCAGCAGTA CGGCCACTAC TTCTGCGAGG CCCAGGAGGG	2820
45	CTCCTACTTC CGCGAGGCTC AGCACTGGCA GCTGCTGCC GAGGACGGCA TCATGGCCGA	2880
	GCACCTGCTG GGTCACTGCC GTGCCCTGGC TGCCTCCCTC TGGCTGGGG TGCTGCCAC	2940
	ACTCACTCTT GGCTTGCTGG TCCACATGGT GAGCAAGGGC GAGGAGCTGT TCACCGGGGT	3000
50	GGTGCCCCATC CTGGTCGAGC TGGACGGCGA CGTAAACGGC CACAAGTTCA GCGTGTCCGG	3060
	CGAGGGCGAG GCGGATGCCA CCTACGGCAA GCTGACCCCTG AAGTTCATCT GCACCACCGG	3120

5	CAAGCTGCC C GTGCCCTGGC CCACCCCTCGT GACCACCCGT ACCTACGGCG TGCAGTGCTT	3180
	CAGCCGCTAC CCCGACCACA TGAAGCAGCA CGACTTCTTC AAGTCCGCCA TGCCCGAAGG	3240
10	CTACGTCCAG GAGCGCACCA TCTTCTTCAA GGACGACGGC AACTACAAGA CCCGGCCCGA	3300
	GGTGAAGTTC GAGGGCGACA CCCTGGTGA CCGCATCGAG CTGAAGGGCA TCGACTTCAA	3360
15	GGAGGACGGC AACATCCTGG GGCACAAAGCT GGAGTACAAC TACAACAGCC ACAACGTCTA	3420
	TATCATGGCC GACAAGCAGA AGAACGGCAT CAAGGTGAAC TTCAAGATCC GCCACAACAT	3480
20	CGAGGACGGC AGCGTGCAGC TCGCCGACCA CTACCAAGCAG AACACCCCCA TCGCGACGG	3540
	CCCCGTGCTG CTGCCCCACA ACCACTACCT GAGCACCCAG TCCGCCCTGA GCAAAGACCC	3600
	CAACGAGAAG CGCGATCACA TGGTCTGCT GGAGTTCTG ACCGCCGCCG GGATCACTCT	3660
25	CGGCATGGAC GAGCTGTACA AGGTGAAGCT TGGGCCCCAA CAAAAACTCA TCTCAGAAGA	3720
	GGATCTGAAT AGCGCCGTCG ACCATCATCA TCATCATCAT TGAGTTAAA CCGCTGATCA	3780
	GCCTCGACTG TGCCCTCTAG TTGCCAGCCA TCTGTTGTTT GCCCCCTCCCC CGTGCCTTCC	3840
30	TTGACCCCTGG AAGGTGCCAC TCCCACGTGTC CTTTCTTAAT AAAATGAGGA AATTGCATCG	3900
	CATTGTCTGA GTAGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA CAGCAAGGGG	3960
	GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCCG TGGGCTCTAT GGCTTCTGAG	4020
35	GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG TATCCCCACG CGCCCTGTAG CGGCGCATTA	4080
	AGCGCGGCCG GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG	4140
	CCCGCTCCCT TCGCTTCTT CCCCTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA	4200
	GCTCTAAATC GGGGCATCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC	4260
40	AAAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT	4320
	CGCCCTTGA CGTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA	4380
	ACACTCAACC CTATCTCGGT CTATCTTTT GATTTATAAG GGATTTGGG GATTTCGGCC	4440
45	TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG	4500
	TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG TATGCAAAGC	4560
	ATGCATCTCA ATTAGTCAGC AACCAAGGTGT GGAAAGTCCC CAGGCTCCCC AGCAGGCAGA	4620
50	AGTATGCAAAG CATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCCT AACTCCGCC	4680
	ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT	4740

5	TTTATTTATG CAGAGGCCGA GGCGCCCTCT GCCTCTGAGC TATTCCAGAA GTAGTGAGGA	4800
	GGCTTTTTG GAGGCCTAGG CTTTGCAAA AAGCTCCCGG GAGCTTGAT ATCCATTTC	4860
	GGATCTGATC AAGAGACAGG ATGAGGATCG TTTCGCATGA TTGAACAAGA TGGATTGCAC	4920
	GCAGGGTTCTC CGGGCCCTTG GGTGGAGAGG CTATTCCGCT ATGACTGGGC ACAACAGACA	4980
10	ATCGGCTGCT CTGATGCCGC CGTGTCCGG CGTCAGCGC AGGGGCGCCC GGTTCTTTT	5040
	GTCAAGACCG ACCTGTCCGG TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG	5100
	TGGCTGGCCA CGACGGCGT TCCTTGCAGA GCTGTGCTCG ACGTTGTCAC TGAAGCGGA	5160
15	AGGGACTGGC TGCTATTGGG CGAAGTGCAGG GGGCAGGATC TCCTGTCAATC TCACCTTGCT	5220
	CCTGCCGAGA AAGTATCCAT CATGGCTGAT GCAATGCAGC GGCTGCATAC GCTTGATCCG	5280
	GCTACCTGCC CATTGACCA CCAAGCGAAA CATCGCATCG AGCGAGCAGC TACTCGGATG	5340
20	GAAGCCGGTC TTGTCGATCA GGATGATCTG GACGAAGAGC ATCAGGGCT CGCGCCAGCC	5400
	GAACTGTTCG CCAGGCTCAA GGCGCGCATG CCCGACGGCG AGGATCTCGT CGTGACCCAT	5460
	GGCGATGCCT GCTTGCAGA TATCATGGTG GAAAATGGCC GCTTTCTGG ATTCAATCGAC	5520
25	TGTGGCCGGC TGGGTGTGGC GGACCGCTAT CAGGACATAG CGTTGGCTAC CGGTGATATT	5580
	GCTGAAGAGC TTGGCGCGA ATGGGCTGAC CGCTTCCCTCG TGCTTTACGG TATCGCCGCT	5640
	CCCGATTCCG AGCGCATCGC CTTCTATCGC CTTCTTGACG AGTTCTCTG AGCGGGACTC	5700
30	TGGGGTTCGA AATGACCGAC CAAGCGACGC CCAACCTGCC ATCACGAGAT TTCAATCCTA	5760
	CCGGCCGCTT CTATGAAAGG TTGGGCTTCG GAATCGTTT CCGGGACGCC GGCTGGATGA	5820
	TCCTCCAGCG CGGGGATCTC ATGCTGGAGT TCTTCGCCA CCCCAACTTG TTTATTGCAG	5880
35	CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATT CACAAATAAA GCATTTTTT	5940
	CACTGCATTC TAGTTGTGGT TTGTCAAAC TCATCAATGT ATCTTATCAT GTCTGTATAC	6000
40	CGTCGACCTC TAGCTAGAGC TTGGCGTAAT CATGGTCATA GCTGTTCCCT GTGTGAAATT	6060
	GTTATCGCT CACAATTCCA CACAACATAC GAGCCGGAAAG CATAAAGTGT AAAGCCTGGG	6120
	GTGCCTAATG AGTGAGCTAA CTCACATTAA TTGCGTTGCG CTCACTGCC GCTTTCCAGT	6180
45	CGGGAAACCT GTCGTGCAG CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT	6240
	TGCGTATTGG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCGGC	6300
50	TGCGGGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG	6360
	ATAACGCAGG AAAGAACATG TGAGCAAAAG GCCAGCAAA GGCCAGGAAC CGTAAAAGG	6420

	CCGCCTTGCT GGCCTTTTC CATAAGGCTCC GCCCCCTGA CGAGCATCAC AAAAATCGAC	6480
5	GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAG ATACCAAGGCG TTTCCCCCTG	6540
	GAAGCTCCCT CGTGCCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCT	6600
	TTCTCCCTTC GGGAAAGCGTG GCGCTTCTC AATGCTCACG CTGTAGGTAT CTCAGTTCGG	6660
10	TGTAGTCGT TCGCTCCAAG CTGGGCTGTG TGCACCAACC CCCCGTTAG CCCGACCGCT	6720
	GCGCCTTATC CGGTAACAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATGCCAC	6780
	TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGT TGTAAGGCGT GCTACAGAGT	6840
15	TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTTGGT ATCTGCGCTC	6900
	TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA	6960
20	CCGCTGGTAG CGGTGGTTTT TTTGTTGCA ACCAGCAGAT TACGCGCAGA AAAAAAGGAT	7020
	CTCAAGAAGA TCCTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACCTAC	7080
	GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT	7140
25	AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC	7200
	AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTGTTCA TCCATAGTTG	7260
	CCTGACTCCC CGTCGTGTAG ATAACCTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG	7320
30	CTGCAATGAT ACCCGGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC	7380
	CAGCCGGAAG GGCGGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC ATCCAGTCTA	7440
	TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTG CGCAACGTTG	7500
35	TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGT TGTTATGGCT TCATTCAAGCT	7560
	CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCAT GTTGTGCAA AAAGCGGTTA	7620
	GCTCCTTCGG TCCTCCGATC GTTGTAGAA GTAAGTTGGC CGCAGTGTAA TCACTCATGG	7680
40	TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA	7740
	CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT CGGGCGACCG AGTTGCTCTT	7800
45	CCCCGGCGTC AATAACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA	7860
	TTGGAAAACG TTCTTCGGGG CGAAAACCTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT	7920
50	CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAAGCATC TTTTACTTTTC ACCAGCGTT	7980
	CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGATAAGG GCGACACCGA	8040

5	AATGTTGAAT ACTCATACTC TTCCCTTTTC AATATTATTG AAGCATTAT CAGGGTTATT	8100
	GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA GGGGTTCCGC	8160
	GCACATITCC CCGAAAAGTG CCACCTGACG TC	8192

(2) INFORMATION ZU SEQ ID NO: 37:

10 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 7000 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

15 (ii) ART DES MOLEKÜLS: DNS (genomisch)

20 (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..7000

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37:

25	AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTCAT TGTTCTCGTT	60
	AGCTCGATGG ACAAGTCAT TGTTCTCTTG CTGAAAGCTC GATGGACAAG TGCATTGTT	120
30	TCTTGCTGAA AGCTCGATGG ACAAGTCAT TGTTCTCTTG CTGAAAGCTC AGTACCCGGG	180
	AGTACCCCTCG ACCGCCGGAG TATAAATAGA GGCCTTCGT CTACGGAGCG ACAATTCAAT	240
	TCAAACAAAGC AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT	300
35	GAACAAGCTA ACAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTA AAAGTAACCA	360
	GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA GAAGTAATTA TTGAATACAA	420
	GAAGAGAACT CTGAATACTT TCAACAAGTT ACCGAGAAAG AAGAAACTCAC ACACAGCTAG	480
40	CGTTTAAACT TAAGCTTGGT ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGGAATTCCG	540
	CTTGGGATGA CGCCTCCCTCC GCCCGGACGT GCCGCCCCCA GCGCACCGCG CGCCCGCGTC	600
45	CCTGGCCCCGC CGGCTCGGTT GGGGCTTCCG CTGCGGCTGC GGCTGCTGCT GCTGCTCTGG	660
	GGGGCCGCCG CCTCCGCCCA GGGCCACCTA AGGAGCCGAC CCCGCATCTT CGCCGTCTGG	720
	AAAGGCCATG TAGGGCAGGA CGGGTGGAC TTTGGCCAGA CTGAGCCGCA CACGGTGCTT	780
50	TTCCACGAGC CAGGCAGCTC CTCTGTGTGG GTGGGAGGAC GTGGCAAGGT CTACCTCTTT	840
	GACTTCCCCG AGGGCAAGAA CGCATCTGTG CGCACGGTGA ATATCGGCTC CACAAAGGGG	900

	TCCGTGCTGG ATAAGCGGGA CTGCGAGAAC TACATCACTC TCCTGGAGAG GCGGAGTGAG	960
5	GGGCTGCTGG CCTGTGGCAC CAACGCCCCG CACCCAGCT GCTGGAACCT GGTGAATGGC	1020
	ACTGTGGTGC CACTTGGCGA GATGAGAGGC TACGCCCCCT TCAGCCCCGGA CGAGAACTCC	1080
	CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT TCCACCATCC GGAAGCAGGA ATACAATGGG	1140
10	AAGATCCCTC GGTTCCGCGG CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC	1200
	ATGCAGAACC CACAGTTCAT CAAAGCCACC ATCGTGCACC AAGACCAGGC TTACGATGAC	1260
	AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT	1320
15	GTGTCCCGTG TGGCCCAGTT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC ACTGTCAGTC	1380
	TCCAAGTGGA ACACCTTCT GAAAGCCATG CTGGTATGCA GTGATGTC CACCAACAAG	1440
	AACTTCAACA GGCTGCAAGA CGTCTTCTG CTCCCTGACC CCAGCGGCCA GTGGAGGGAC	1500
20	ACCAGGGTCT ATGGTGTTTT CTCCAACCCC TGAACTACT CAGCCGTCTG TGTGTATTCC	1560
	CTCGGTGACA TTGACAAGGT CTTCCGTACC TCCTCACTCA AGGGCTACCA CTCAAGCCTT	1620
	CCCAACCCGC GGCCTGGCAA GTGCCTCCCA GACCAGCAGC CGATACCCAC AGAGACCTTC	1680
25	CAGGTGGCTG ACCGTCACCC AGAGGTGGCG CAGAGGGTGG AGCCCATGGG GCCTCTGAAG	1740
	ACGCCATTGT TCCACTCTAA ATACCACTAC CAGAAAGTGG CCGTTCACCG CATGCAAGCC	1800
	AGCCACGGGG AGACCTTCA TGTGCTTAC CTAACACTACG ACAGGGGCAC TATCCACAAG	1860
30	GTGGTGGAAC CGGGGGAGCA GGAGCACAGC TTGCGCTTCA ACATCATGGA GATCCAGCCC	1920
	TTCCGCCCG CGGCTCCAT CCAGACCATG TCGCTGGATG CTGAGCGGAG GAAGCTGTAT	1980
35	GTGAGCTCCC AGTGGGAGGT GAGCCAGGTG CCCCTGGACC TGTGTGAGGT CTATGGCGGG	2040
	GGCTGCCACG GTTGCCCAT GTCCCGAGAC CCCTACTGCG GCTGGGACCA GGGCCGCTGC	2100
	ATCTCCATCT ACAGCTCCGA ACGGTCACTG CTGCAATCCA TTAATCCAGC CGAGCCACAC	2160
40	AAGGAGTGTC CCAACCCAA ACCAGACAAG GCCCCACTGC AGAAGGTTTC CCTGGCCCCA	2220
	AACTCTCGCT ACTACCTGAG CTGCCCCATG GAATCCGCC ACGCCACCTA CTCATGGCGC	2280
	CACAAGGAGA ACGTGGAGCA GAGCTGCGAA CCTGGTCACC AGAGCCCCAA CTGCATCCTG	2340
45	TTCATCGAGA ACCTCACGGC GCAGCAGTAC GCCCACTACT TCTGCGAGGC CCAGGAGGGC	2400
	TCCTACTTCC GCGAGGCTCA GCACTGGCAG CTGCTGCCCG AGGACGGCAT CATGGCCGAG	2460
50	CACCTGCTGG GTCATGCGCTG TGCCCTGGCT GCCTCCCTCT GGCTGGGGGT GCTGCCACCA	2520
	CTCACTCTTG GCTTGCTGGT CCACGTGAAG CTTGGGCCCG TTTAAACCCG CTGATCAGCC	2580

EP 0 892 047 A2

	TCGACTGTGC CTTCTAGTTG CCAGCCATCT GTTGTGCGC CCTCCCCGT GCCTTCCTTG	2640
5	ACCCCTGGAAG GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT	2700
	TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGTGG GGCAGGACAG CAAGGGGGAG	2760
	GATTGGGAAG ACAATAGCAG GCATGCTGGG GATGCGGTGG CCTCTATGGC TTCTGAGGCG	2820
10	GAAAGAACCA GCTGGGCTC TAGGGGTAT CCCCACGCGC CCTGTAGCGG CGCATTAAGC	2880
	GCGGCGGGTG TGGTGGTAC GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCCCCC	2940
	GCTCCTTCG CTTCTTCCC TTCTTCTC GCCACGTTCG CCGGCTTCC CCGTCAAGCT	3000
15	CTAAATCGGG GCATCCCTT AGGGTCCGA TTTAGTGCTT TACGGCACCT CGACCCAAA	3060
	AAACTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC	3120
	CCTTTGACGT TGGAGTCCAC GTTCTTAAT AGTGGACTCT TGTTCCAAC TGGAAACAACA	3180
20	CTCAACCTA TCTCGGTCTA TTCTTTGAT TTATAAGGA TTTTGGGAT TTGGCCTAT	3240
	TGGTTAAAAA ATGAGCTGAT TTAACAAAAA TTAACCGGA ATTAATTCTG TGGAAATGTGT	3300
25	GTCAGTTAGG GTGTGGAAAG TCCCCAGGCT CCCCAGGCAG GCAGAAGTAT GCAAAGCATG	3360
	CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT	3420
	ATGCAAAGCA TGCATCTAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCCGCCATC	3480
30	CCGCCCCCTAA CTCCGCCAG TTCCGCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTT	3540
	ATTTATGCAG AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC	3600
35	TTTTTGGAG GCCTAGGCTT TTGCAAAAG CTCCCGGGAG CTTGTATATC CATTTCGGA	3660
	TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG ACAAGATGG ATTGCACGCA	3720
	GGTTCTCCGG CCGCTGGGT GGAGAGGCTA TTCCGCTATG ACTGGGCACA ACAGACAATC	3780
40	GGCTGCTCTG ATGCCGCCGT GTTCCGGCTG TCAGCGCAGG GGCGCCGGT TCTTTTGTC	3840
	AAGACCGACC TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCCCG GCTATCGTGG	3900
45	CTGGCCACGA CGGGCGTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA AGCGGGAAAGG	3960
	GACTGGCTGC TATTGGCGA AGTGCCGGGG CAGGATCTCC TGTCATCTCA CCTTGCTCCT	4020
	GCCGAGAAAG TATCCATCAT GGCTGATGCA ATGCCGCCGC TGCATACGCT TGATCCGGCT	4080
50	ACCTGCCCAT TCGACCACCA AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA	4140
	GCCGGTCTTG TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCCGAA	4200

	CTGTTGCCA GGCTCAAGGC GCGCATGCC GACGGCGAGG ATCTCGTCGT GACCCATGGC	4260
5	GATGCCCTGCT TGCGGAATAT CATGGTGGAA AATGGCCGCT TTTCTGGATT CATCGACTGT	4320
	GGCCGGCTGG GTGTGGCGGA CCGCTATCAG GACATAGCGT TGGCTACCCG TGATATTGCT	4380
	GAAGAGCTTG GCGGCGAATG GGCTGACCGC TTCTCGTGC TTACGGTAT CGCCGCTCCC	4440
10	GATTCGCAGC GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG	4500
	GGTTGAAAT GACCGACCAA GCGACGCCA ACCTGCCATC ACGAGATTTC GATTCCACCG	4560
	CCGCCTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTCCG GGACGCCGGC TGGATGATCC	4620
15	TCCAGCGCGG GGATCTCATG CTGGAGTTCT TCGCCCACCC CAACTGTTT ATTGCAGCTT	4680
	ATAATGGTTA CAAATAAACG AATAGCATCA CAAATTTCAC AAATAAAGCA TTTTTTTCAC	4740
	TGCATTCTAG TTGTGGTTG TCCAAACTCA TCAATGTATC TTATCATGTC TGTATACCGT	4800
20	CGACCTCTAG CTAGAGCTTG GCGTAATCAT GGTCTAGCT GTTTCCTGTG TGAAATTGTT	4860
	ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG	4920
	CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCCTC ACTGCCGCT TTCCAGTCGG	4980
25	GAAACCTGTC GTGCCAGCTG CATTAAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC	5040
	GTATTGGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC	5100
	GGCGAGCGGT ATCAGCTCAC TCAAAGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA	5160
30	ACCGAGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG	5220
	CGTTGCTGGC GTTTTCCAT AGGCTCCGG CCCCTGACGA GCATCACAAA AATCGACCGCT	5280
35	CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGCGTTT CCCCCTGGAA	5340
	GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC	5400
	TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC AGTTGGTGT	5460
40	AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG	5520
	CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGTAAG ACACGACTTA TCGCCACTGG	5580
45	CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT	5640
	TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCCTCTGC	5700
	TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTTTG ATCCGGCAA CAAACCACCG	5760
50	CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC	5820
	AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT	5880

	AAGGGATTTC GGTGATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA	5940
5	AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT	6000
	GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT	6060
	GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG	6120
10	CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAAGCCAG	6180
	CCGGAAGGGC CGAGCCGAGA AGTGGTCCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA	6240
	ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG	6300
15	CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTGG TATGGCTTCA TTCAGCTCCG	6360
	GTTCCCAACG ATCAAGGCAGA GTTACATGAT CCCCCATGTT GTGAAAAAA GCGGTTAGCT	6420
20	CCTTCGGTCC TCCGATCGTT GTCAGAAGTA AGTTGGCCGC AGTGTATATCA CTCATGGTTA	6480
	TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG	6540
	GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCC GCGACCGAGT TGCTCTGCC	6600
25	CGGCGTCAAT ACGGGATAAT ACCGGGCCAC ATAGCAGAAC TTTAAAGTG CTCATCATTG	6660
	GAAAACGTTTC TTGGGGCGA AAACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA	6720
	TGTAACCCAC TCGTGCACCC AACTGATCTT CAGCATCTT TACTTTCACC AGCGTTCTG	6780
30	GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAAT	6840
	GTTGAATACT CATACTCTTC CTTTTCAAT ATTATTGAAG CATTATTCAG GGTTATTGTC	6900
35	TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAAATAGGG GTTCCCGCGCA	6960
	CATTTCCCCG AAAAGTGCCA CCTGACGTCG ACGGATCGGG	7000

40 (2) INFORMATION ZU SEQ ID NO: 38:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 7108 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: DNS (genomisch)

50 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..7108

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 38:

5	AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT TGTTCTCGTT	60
10	AGCTCGATGG ACAAGTGCAT TGTTCTCTG CTGAAAGCTC GATGGACAAG TGCATTGTTG TCTTGTGAA AGCTCGATGG ACAAGTGCAT TGTTCTCTG CTGAAAGCTC AGTACCCGGG	120
15	AGTACCCCTCG ACCGCCGGAG TATAATAGA GGCCTTCGT CTACGGAGCG ACAATTCAAT TCAAACAAGC AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT	180
20	GAACAGCTA AACAACTCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTAA AAAGTAACCA GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA GAAGTAATTAA TTGAATACAA	240
25	GAAGAGAACT CTGAATACTT TCAACAAGTT ACCGAGAAAG AAGAARCTCAC ACACAGCTAG CGTTTAAACT TAAGCTTGGT ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGGAATTGG	300
30	CTTGGGATGA CGCCTCCTCC GCCCCGGACGT GCGCCCCCGA GCGCACCGCG CGCCCGCGTC CCTGGCCCGC CGGCTCGGTT GGGGCTTCCG CTGCGCTGCG GGCTGCTGCT GCTGCTCTGG	360
35	GCGGCCGCG CCTCCGCCA GGGCCACCTA AGGAGCGGAC CCCGCATCTT CGCCGTCTGG AAAGGCCATG TAGGGCAGGA CCGGGTGGAC TTTGGCCAGA CTGAGCCGCA CACGGTGCCTT	420
40	TTCCACGAGC CAGGCAGCTC CTCTGTGTGG GTGGGAGGAC GTGGCAAGGT CTACCTCTTT GACTTCCCCG AGGGCAAGAA CGCATCTGTG CGCACGGTGA ATATCGGCTC CACAAAGGGG	480
45	TCCCTGTCTGG ATAAGCGGGA CTGCGAGAAC TACATCACTC TCCTGGAGAG GCGGAGTGTAG GGGCTGCTGG CCTGTGGCAC CAACGCCGG CACCCAGCT GCTGGAACCT GGTGAATGGC	540
50	ACTGTGGTGC CACTTGGCGA GATGAGAGGC TACGCCCT TCAGCCCGA CGAGAACTCC CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT TCCACCATCC GGAAGCAGGA ATACAATGGG AAGATCCCTC GGTTCCCGG CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC	600
	ATGCAGAACCC CACAGTTCAT CAAAGCCACC ATCGTGCACC AAGACCAGGC TTACGATGAC AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT	660
	GTGTCCCGTG TGGCCCAGTT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC ACTGTCAAGTC TCCAAGTGGAA ACACCTTCT GAAAGCCATG CTGGTATGCA GTGATGCTGC CACCAACAAG	720
	AACTTCAACA GGCTGCAAGA CGTCTCCTG CTCCCTGACC CCAGCGGCCA GTGGAGGGAC ACCAGGGTCT ATGGTGTGTTT CTCCAACCCCC TGAAACTACT CAGCCGTCTG TGTGTATTCC	780
		1140
		1200
		1260
		1320
		1380
		1440
		1500
		1560

	CTCGGTGACA TTGACAAGGT CTTCCGTACC TCCTCACTCA AGGGCTACCA CTCAAGCCTT	1620
5	CCCAACCCGC GGCCTGGCAA GTGCCTCCCA GACCAGCAGC CGATACCCAC AGAGACCTTC	1680
	CAGGTGGCTG ACCGTACCCC AGAGGTGGCG CAGAGGTGG AGCCCATGGG GCCTCTGAAG	1740
	ACGCCATTTGT TCCACTCTAA ATACCACTAC CAGAAAGTGG CCGTTCACCG CATGCAAGCC	1800
10	AGCCACGGGG AGACCTTCA TGTGCTTAC CTAACTACAG ACAGGGGCAC TATCCACAAG	1860
	GTGGTGGAAAC CGGGGGAGCA GGAGCACAGC TTGCGCTTCA ACATCATGGA GATCCAGCCC	1920
	TTCCGCGCG CGGCTGCCAT CCAGACCATG TCGCTGGATG CTGAGCGGAG GAAGCTGTAT	1980
15	GTGAGCTCCC AGTGGGAGGT GAGCCAGGTG CCCCTGGACC TGTGTGAGGT CTATGGCGGG	2040
	GGCTGCCACG GTTGCTCAT GTCCCGAGAC CCCTACTGCG GCTGGGACCA GGGCCGCTGC	2100
	ATCTCCATCT ACAGCTCCGA ACGGTCAGTG CTGCAATCCA TTAATCCAGC CGAGCCACAC	2160
20	AAGGAGTGTC CCAACCCCAA ACCAGACAAG GCCCCACTGTC AGAAGGTTTC CCTGGCCCCA	2220
	AACTCTCGCT ACTACCTGAG CTGCCCCATG GAATCCCGCC ACGCCACCTA CTCATGGCGC	2280
	CACAAGGAGA ACGTGGAGCA GAGCTGCGAA CCTGGTCACC AGAGCCCCAA CTGCATCCTG	2340
25	TTCATCGAGA ACCTCACGGC GCAGCAGTAC GGCCACTACT TCTGCGAGGC CCAGGAGGGC	2400
	TCCTACTTCC GCGAGGCTCA GCACTGGCAG CTGCTGCCG AGGACGGCAT CATGGCCGAG	2460
	CACCTGCTGG GTCATGCCCTG TGCCCTGGCT GCCTCCCTCT GGCTGGGGGT GCTGCCACAA	2520
30	CTCACTCTTG GCTTGTGGT CCACGTGAAG CTTGGGCCG AACAAAAACT CATCTCAGAA	2580
	GAGGATCTGA ATAGGCCGT CGACCATCAT CATCATCATC ATTGAGTTA TCCAGCACAG	2640
	TGGCGGCCGC TCGAGTCTAG AGGGCCCGTT TAAACCCGCT GATCAGCCTC GACTGTGCCT	2700
35	TCTAGTTGCC AGCCATCTGT TGTGCCCC TCCCCCGTGC CTTCTTGAC CCTGGAAGGT	2760
	GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAATTG CATCGCATTG TCTGAGTAGG	2820
40	TGTCATTCTA TTCTGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA TTGGGAAGAC	2880
	AATAGCAGGC ATGCTGGGA TGCGGTGGGC TCTATGGCTT CTGAGGGGA AAGAACCCAGC	2940
45	TGGGGCTCTA GGGGGTATCC CCACCGGCCG TGTAGCGGGG CATTAAAGCGC GGCGGGTGTG	3000
	GTGGTTACGC GCACCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT	3060
	TTCTTCCCTT CCTTTCTCGC CACGTTCGCC GGCTTCCCC GTCAAGCTCT AAATCGGGGC	3120
50	ATCCCTTCTAG GGTTCCGATT TAGTGCTTCA CGGCACCTCG ACCCCAAAAA ACTTGATTAG	3180
	GGTGATGGTT CACGTAGTGG GCCATCGCCC TGATAGACGG TTTTCGCC TTTGACGTTG	3240

EP 0 892 047 A2

	GAGTCCACGT TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC	3300
5	TCGGTCTATT CTTTGATTT ATAAGGGATT TTGGGGATTT CGGCCTATTG GTTAAAAAAT	3360
	GAGCTGATTT AACAAAAATT TAACCGAAT TAATCTGTG GAATGTGTGT CAGTTAGGGT	3420
	GTGAAAAGTC CCCAGGCTCC CCAGGCAGGC AGAAGTATGC AAAGCATGCA TCTCAATTAG	3480
10	TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAAGTAT GCAAAGCATG	3540
	CATCTCAATT AGTCAGCAAC CATACTCCCG CCCCTAACTC CGCCCATCCC GCCCCTAACT	3600
	CCGCCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTTAT TTATGCAGAG	3660
15	GCCGAGGCCG CCTCTGCCTC TGAGCTATT CAGAAGTAGT GAGGAGGCTT TTTTGGAGGC	3720
	CTAGGCTTT GCAAAAAGCT CCCGGGAGCT TGTATATCCA TTTTCGGATC TGATCAAGAG	3780
	ACAGGATGAG GATCGTTTCG CATGATTGAA CAAGATGGAT TGCACCGAGG TTCTCCGGCC	3840
20	GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAAC AGACAATCGG CTGCTCTGAT	3900
	GCCGCCGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTT TCAGGCTT GACCGACCTG	3960
25	TCCGGTGCC C TGAATGAACT GCAGGACGAG GCAGCGCGGC TATCGTGGCT GGCACGACG	4020
	GGCGTTCCCT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAAGGA CTGGCTGCTA	4080
	TTGGCGAAG TGCCGGGCA GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA	4140
30	TCCATCATGG CTGATGCAAT GCGCGGCTG CATACTGTTG ATCCGGCTAC CTGCCCATTC	4200
	GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC CGGTCTTGT	4260
	GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC CAGCCGAACG GTTCGCCAGG	4320
35	CTCAAGGCAGC GCATGCCGA CGGGAGGAT CTCGCTGTGA CCCATGGCGA TGCCTGCTT	4380
	CCGAATATCA TGGTGGAAA TGGCGCTTT TCTGGATTCA TCGACTGTGG CGGGCTGGGT	4440
40	GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA AGAGCTTGGC	4500
	GGCGAATGGG CTGACCGCTT CCTCGTGTCT TACGGTATCG CCGCTCCCGA TTCGCAGCGC	4560
	ATCGCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG GACTCTGGGG TTCGAATGA	4620
45	CCGACCAAGC GACGCCAAC CTGCCATCAC GAGATTCGA TTCCACCGCC GCCTTCTATG	4680
	AAAGGTTGGG CTTCGGAATC GTTTCCGGG ACGCCGGCTG GATGATCCTC CAGCGCGGG	4740
50	ATCTCATGCT GGAGTTCTTC GCCCACCCCA ACTTGTTTAT TGCAGCTTAT AATGGTTACA	4800
	AATAAAAGCAA TAGCATCACA AATTTCACAA ATAAAGCATT TTTTTCACTG CATTCTAGTT	4860

	GTGGTTTGTCAAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCT	4920
5	AGAGCTTGGCTAATCATGGTCATAGCTGTTCTGTGTGAAATIGTTATCCGCTCACAA	4980
	TTCCACACAACTACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGTGCTTAATGAGTGA	5040
	GCTAACTCACATTAAATTGCGTTGCGCTCAC TGCCCCGTTTCCAGTCGGGA AACCTGTCGT	5100
10	GCCAGCTGCTTAATGAATC GGCCAACGCGCGGGGAGAGGCGGTTGCGTATTGGCGCT	5160
	CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGTTCGGCTGCGCGAGCGGTAT	5220
	CAGCTCACTC AAAGGCCGTAATACGGTTATCCACAGAATCAGGGATAACGCAGGAAAGA	5280
15	ACATGTGAGC AAAAGGCCAG CAAAAGGCCAGAACCCGTTGCTGGCGTTTCCAGGCG	5340
	TTTTCATAGGCTCCGCCCCCCTGACGAGC ATCACAAAAATCGACGCTCAAGTCAGAGGT	5400
	GGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCCCTGGAAGCTCCCTCGTGC	5460
20	GCTCTCTGTCTCGACCCCTGCGCTTACCGGATACCTGTC CGCCCTTCTCCTTCGGGAA	5520
	GCGTGGCGCTTTCTCAATGCTACGCTGTA GGTATCTCAGTTCGGTGAGGTCGTTCGCT	5580
	CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTAGCCGAGCCCTATCCGTA	5640
25	ACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCAGCCACTGGCA GCAGCCACTG	5700
	GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGCTACAGAGTTCTTGAAAGGGTGGC	5760
30	CTAACTACGGCTACACTAGAAGGACAGTATTTGCTGCGCTGCTGAAAGCCAGTTA	5820
	CCTTCGGAAAAGAGTTGGTACCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTG	5880
	GTTCGGGTTGTTGAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAAGATCCTT	5940
35	TGATCTTTCTACGGGTCTGACGCTCAGTGGAAACGAACTCAGTTAAAGGATTTGG	6000
	TCATGAGATTATCAAAAGGATCTCACCTAGATCCTTTAAATTTAAAGTTTTA	6060
	AATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG	6120
40	AGGCACCTATCTCAGCGATCTGCTATTCTGTCATCCATAGTGCCTGACCTGGCTCG	6180
	TGTAGATAAC TACGATAACGGGAGGCTTACCATCTGGCCCATGCTGCAATGATAACCGC	6240
45	GAGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAAACAGCCAGCCGGAAAGGGCG	6300
	AGCGCAGAAGTGGCTCTGCAACTTATCCGCCTCCATCCA GTCTATTAAATGTTGCCGGG	6360
	AAGCTAGAGTAAAGTAGTTCGCCAGTTAATAAGTTGCGCAA CGTTGTTGCCATTGCTACAG	6420
50	GCATCGTGGTGTACACGCTCGTCGTTGGTATGGCTTCATTCAAGCTCCGGTTCCAAACGAT	6480
	CAAGGGAGTTACATGATCCCCATGTTGTGCAAAAGC GGTTAGCTCCGGGTCTC	6540

5	CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT CATGGTTATG GCAGCACTGC	6600
	ATAATTCTCT TACTGTCATG CCATCCGTAAT GATGCTTTTC TGACTGGGT GAGTACTCAA	6660
	CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCCG GCGTCAATAC	6720
10	GGGATAATAC CGGGCCACAT AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT	6780
	CGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC CAGTCGATG TAACCCACTC	6840
	GTGCACCCAA CTGATCTTCA GCATCTTTA CTTTCACCAAG CGTTTCTGGG TGAGCAAAA	6900
15	CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA	6960
	TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT	7020
	ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGT TCCGGGCACA TTTCCCCGAA	7080
20	AAGTGCCACC TGACGTCGAC GGATCGGG	7108

(2) INFORMATION ZU SEQ ID NO: 39:

- (i) SEQUENZ CHARAKTERISTIKA:
 - (A) LÄNGE: 4019 Basenpaare
 - (B) ART: Nukleinsäure
 - (C) STRANGFORM: Einzel
 - (D) TOPOLOGIE: linear
- (ii) ART DES MOLEKÜLS: DNS (genomisch)
- 30 (ix) MERKMALE:
 - (A) NAME/SCHLÜSSEL: exon
 - (B) LAGE: 1..4019

35	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 39:	
	CTCGAGAAAT CATAAAAAT TTATTGCTT TGAGCGGA TAACAATTAT AATAGATTCA	60
40	ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGAGA	120
	GGATCCATC ACCATCACCA TCACGGATCC CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT	180
	TCCACCATCC GGAAGCAGGA ATACAATGGG AAGATCCCTC GGTTCCGCCG CATCCGGGC	240
45	GAGAGTGAGC TGACACCAAG TGATACTGTC ATGCAGAACCC CACAGTTCAT CAAAGCCACC	300
	ATCGTGCACC AAGACCAGGC TTACGATGAC AAGATCTACT ACTTCTTCG AGAGGACAAT	360
50	CCTGACAAGA ATCCTGAGGC TCCTCTCAAT GTGTCGGTG TGGCCCGTT GTGCAGGGGG	420
	GACCAGGGTG GGGAAAGTTC ACTGTCAGTC TCCAAGTGGAA ACACTTTCT GAAAGCCATG	480

	CTGGTATGCA GTGATGCTGC CACCAACAAG AACTTCAACA GGCTGCAAGA CGTCTTCCTG	540
5	CTCCCTGACC CCAGCGGCCA GTGGAGGGAC ACCAGGGTCT ATGGTGTITCTCCAAACCCC	600
	TGGAACTACT CAGCCGTCTG TGTGTATTCC CTCGGTGACA TTGACAAGGT CTTCCGTACC	660
	TCCTCACTCA AGGGCTACCA CTCAAGCCTT CCCAACCCGC GCCCTGGCAA GTGCCTCCCA	720
10	GACCAGCAGC CGATACCCAC AGAAAGCTTA ATTAGCTGAG CTTGGACTCC TGTTGATAGA	780
	TCCAGTAATG ACCTCAGAAC TCCATCTGGA TTTGTTCAGA ACGCTCGGTT GCCGCCGGC	840
	GTTTTTTATT GGTGAGAATC CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA	900
15	AAATGGAGAA AAAATCACT GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG	960
	AACATTTGAGGCATTTCAAG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG	1020
20	ATATTACGGC CTTTTTAAAG ACCGTAAGA AAAATAAGCA CAAGTTTAT CCGGCCTITA	1080
	TTCACATTCT TGCCCCCTG ATGAATGCTC ATCCGGAATT TCGTATGGCA ATGAAAGACG	1140
	GTGAGCTGGT GATATGGGAT AGTGTTCACC CTTGTTACAC CGTTTCCAT GAGCAAACGT	1200
25	AAACGTTTTC ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT	1260
	ATTCGCAAGA TGTGGCGTGT TACGGTGAAA ACCTGGCTA TTTCCCTAAA GGTTTATTG	1320
	AGAATATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTT GATTTAAACG	1380
30	TGGCCAATAT GGACAACCTTC TTGCCCCCG TTTTCACCAT GGGCAAATAT TATACGCAAG	1440
	GCGACAAGGT GCTGATGCCG CTGGCGATTCA AGGTCATCA TGCCGCTGT GATGGCTTCC	1500
35	ATGTCGGCAG AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGCGT	1560
	AATTTTTTA AGGCAGTTAT TGGTGCCTT AAACGCTGG GGTAATGACT CTCTAGCTTG	1620
	AGGCATCAA TAAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTTGTTT TATCTGTTGT	1680
40	TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCG TCTAGAGCTG CCTCGCGCGT	1740
	TTCGGTGATG ACGGTGAAAA CCTCTGACAC ATGCAAGCTCC CGGAGACGGT CACAGCTTGT	1800
	CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG	1860
45	TGTCGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATACT TGGCTTAACT	1920
	ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT GCGGTGTGAA ATACCGCACA	1980
50	GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTCCGC TTCCCTCGCTC ACTGACTCGC	2040
	TGCGCTCGGT CTGTCGGCTG CGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT	2100

	TATCCACAGA ATCAGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG	2160
5	CCAGGAACCG TAAAAAAGGC GCGITGCTGG CGTTTTCCA TAGGCTCCGC CCCCCTGACG	2220
	AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT	2280
10	ACCAGGGCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA	2340
	CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTCTCAA TGCTCACGCT	2400
15	GTAGGTATCT CAGTCGGTG TAGGTCGTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC	2460
	CCGTTCAAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA	2520
20	GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG	2580
	TAGGCCTGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGGACAG	2640
25	TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT	2700
	GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTG TGTTTCAAG CAGCAGATT	2760
	CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC	2820
30	AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA	2880
	CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA	2940
	CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT	3000
35	TTCTGTTACATC CATAGCTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT	3060
	TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT	3120
	TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT	3180
40	CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT TCGCCAGTTA	3240
	ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG	3300
	GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT	3360
45	TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAAGT AAGTTGGCCG	3420
	CAGTGTATTC ACTCATGGTT ATGGCAGCAC TGCTATAATTC TCTTACTGTC ATGCCATCCG	3480
	TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC	3540
50	GGCGACCGAG TTGCTCTTGC CCGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA	3600
	CTTTAAAAAGT GCTCATCATT GGAAAACGTT CTTGGGGCG AAAACTCTCA AGGATCTTAC	3660
	CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT	3720
	TTACTTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG	3780

5 GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA 3840
 GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA 3900
 AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGC ACCTGACGTC TAAGAAACCA 3960
 TTATTATCAT GACATTAACC TATAAAAATA GCGTATCAC GAGGCCCTTT CGTCTTCAC 4019

10 (2) INFORMATION ZU SEQ ID NO: 40:

15 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 3999 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

20 (ii) ART DES MOLEKÜLS: DNS (genomisch)

25 (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..3999

25 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 40:

30 CTCGAGAAAT CATAAAAAT TTATTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA 60
 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGAGA 120
 GGATCGCATC ACCATCACCA TCACACGGAT CGGCATGCGA GCTCCAGTG GGAGGTGAGC 180
 CAGGTGCCCC TGGACCTGTG TGAGGTCTAT GGCGGGGCT GCCACGGTTG CCTCATGTCC 240
 35 CGAGACCCCT ACTGCGGCTG GGACCAGGGC CGCTGCATCT CCATCTACAG CTCCGAACGG 300
 TCAGTGCTGC AATCCATTAA TCCAGCCGAG CCACACAAGG AGTGTCCCAA CCCAAACCA 360
 GACAAGGCC CACTGCAGAA GGTTCCCTG GCCCCAAACT CTCGCTACTA CCTGAGCTGC 420
 40 CCCATGGAAT CCCGCCACGC CACCTACTCA TGGGCCACA AGGAGAACGT GGAGCAGAGC 480
 TGCGAACCTG GTCACCAGAG CCCCAACTGC ATCCCTGTTCA TCGAGAACCT CACGGCGCAG 540
 45 CAGTACGGCC ACTACTTCTG CGAGGCCAG GAGGGCTCCT ACTTCCGCGA GGCTCAGCAC 600
 TGGCAGCTGC TGCCCGAGGA CGGCATCATG GCCGAGCACC TGCTGGTCA TGCCTGTGCC 660
 CTGGCTGCCT CCCTCTGGCT GGGGGTGCTG CCCACACTCA CTCTTGGCTT GCTGGTCCAC 720
 50 GTGAAGCTTA ATTAGCTGAG CTTGGACTCC TGTGATAGA TCCAGTAATG ACCTCAGAAC 780
 TCCATCTGGA TTTGTTAGA ACCGCTCGGTT GCCGCCGGGC GTTTTTTATT GGTGAGAAC 840

EP 0 892 047 A2

	CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT	900
5	GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTGA GGCATTTCA	960
	TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTTAAAG	1020
	ACCGTAAAGA AAAATAAGCA CAAGTMTTAT CCGGCCTTTA TTCACTATTCT TGCCCGCCTG	1080
10	ATGAATGCTC ATCCGGAATT TCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT	1140
	AGTGTTCACC CTGTTACAC CGTTTCCAT GAGCAAACGT AAACGTTTC ATCGCTCTGG	1200
	AGTGAATACC ACCGACGATT CCGGCAGTTT CTACACATAT ATTGCAAGA TGTGGCGTGT	1260
15	TACGGTGAAA ACCTGGCTA TTTCCTAAA GGGTTATTG AGAATATGTT TTTCGTCTCA	1320
	GCCAATCCCT GGGTGAGTTT CACCAGTTT GATTTAACG TGGCAATAT GGACAACCTTC	1380
20	TTCCGCCCCG TTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG	1440
	CTGGCGATTG AGGTTCATCA TGCCGCTGT GATGGCTTC ATGCGGCAG AATGCTTAAT	1500
	GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGCGT AATTTTTTA AGGCAGTTAT	1560
25	TGGTGCCTT AAACGCCCTGG GGTAATGACT CTCTAGCTTG AGGCATCAAA TAAAACGAAA	1620
	GGCTCAGTCG AAAGACTGGG CCTTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT	1680
	GAGTAGGACA AATCCGCCCC TCTAGAGCTG CCTCGCGCGT TTGGTGATG ACGGTGAAAA	1740
30	CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTGT CTGTAAGCGG ATGCCGGGAG	1800
	CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG TGTCGGGGCG CAGCCATGAC	1860
35	CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACG ATGCGGCATC AGAGCAGATT	1920
	GTACTGAGAG TGCACCATAT GCGGTGTGAA ATACCGCACA GATGCGTAAG GAGAAAATAC	1980
	CGCATCAGGC GCTCTTCCGC TTCTCGCTC ACTGACTCGC TGGCCTCGGT CTGTCGGCTG	2040
40	CGGGGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT	2100
	AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAGGCC	2160
45	GCGTTGCTGG CGTTTTCCA TAGGCTCCGC CCCCCGTACG AGCATCACAA AAATCGACGC	2220
	TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAGAT ACCAGGGGTT TCCCCCTGGA	2280
	AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGGCTT	2340
50	CTCCCTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT GTAGGTATCT CAGTTGGTG	2400
	TAGGTCGTTG GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC	2460

	GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG	2520
5	GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCAGTGC TACAGAGTTC	2580
	TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG	2640
	CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGAA ACAAAACCACC	2700
10	GCTGGTAGCG GTGGTTTTTG TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT	2760
	CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT	2820
	TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTCA CCTAGATCCT TTTAAATTAA	2880
15	AAATGAAGTT TAAATCAAT CTAAAGTATA TATGAGTAA CTTGGTCTGA CAGTTACCAA	2940
	TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCTGTCATC CATACTGCC	3000
20	TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCCAGTGCT	3060
	GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAAGCCA	3120
	GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT	3180
25	AATTGTTGCC GGGAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT	3240
	GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCACTGCC	3300
	GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGAAAAAA AGCGGTTAGC	3360
30	TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT	3420
	ATGGCAGCAC TGCTAAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT	3480
35	GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC	3540
	CCGGCGTCAA TACGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAGT GCTCATCATT	3600
	GGAAAACGTT CTTGGGGCG AAAACTCTCA AGGATCTAC CGCTGTTGAG ATCCAGTTCG	3660
40	ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTCT	3720
	GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAGG GAATAAGGGC GACACGGAAA	3780
45	TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT	3840
	CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC	3900
	ACATTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC	3960
50	TATAAAAATA GGCGTATCAC GAGGCCCTT CGTCTTCAC	3999

(2) ANGABEN ZU SEQ ID NO: 41:

5 (i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 8888 Basenpaare
 (B) ART: Nucleotid
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Genom-DNA

15 (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..8888

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 41:

20	GAGCCGCACA CGGTGCTTT CCACGAGCCA GGCAGCTCCT CTGTGTGGT GGGAGGACGT	60
25	GGCAAGGTCT ACCTCTTGA CTTCCCCGAG GGCAAGAACG CATCTGTGCG CACGGTGAC	120
30	CTCTCTCTTC CCCCCAACACC CCCCCTACCC TCTTATCTCC CCTCTGGCCC TGCCAAGGGT	180
35	CCTCAGGGAA TCCGAGGGAG CTGGCTTCTC TTCCTAACT GCCCCCACCT CCGTATCCTA	240
40	TAAATGGCTC CTGGGGGAGG CTCCCTAAAG GTAGTCCAGA TTGGACTGGG GAGCTGGGC	300
45	GGTGTGGAGA AAAACAGGGAG CTAATGGGCC TGGCCAGCTG GGCAGGGCTG CTGCGGAAAG	360
50	CCCAGGCTGG AAGCTGGGCC CCAGAGCCCA TGCCTGGTCT TCTGAACCCCT CTGGGCCTCA	420
55	GCTCTGGATA TGAGACCCCTG TTTGACCTCA GGTAGATCAC TCACCCCTCTC AGAGCCCCAG	480
60	TTGCTCATCT GTCAAGATGAG AATAATGGTT GCTTCCTTGT GGGCTTATCC TGAGGCTGTG	540
65	TGGAAAGCAT TTCAGGGGTA CCTCACCCCT GGCAGATTGA ACTAATGCTT CTCCCTTCC	600
70	CCAGGGTAAT ATCGGCTCCA CAAAGGGGTC CTGTCTGGAT AAGCGGGTGA CGGGGGGAGG	660
75	GATCTGGAGG GGTCTGAGCC ACTTGGTAA GGGAGAGGGAG ACCCTGAGGG TCTAAGGAAG	720
80	GAAGCATGGC CCTGCCCCAC GAGTCCCAGA CTGATGGGA GACGTGGTCC TCTGTGCTTA	780
85	GGGGATGGCG TCAGCTGCAC ACACTCTGGG CTGTCCCCGG AGGCTGTCAC CTATGCTAAG	840
90	CCCTCTGAC ACCTCTTCC CTGATCTGG GGGTCTAGT GCTAGGCTTG CCAGGGCTT	900
95	CCAGCAACCA ATTTCTCTCC TCCCTTCTCT CTTCCCCGGG CAGGACTGCG AGAACTACAT	960
100	CACTCTCCTG GAGAGGCGGA GTGAGGGGCT GCTGGCTGT GGCACCAACG CCCGGCACCC	1020
105	CAGCTGCTGG AACCTGGTGA GAAGGCTGCT CCCCATGTGC CTGATCAGCT CACCTTCTAC	1080
110	TGCGTGGGCT TCTGCCCCCTC ATGGTGGGAA GGAGATGGCG AGACTCCAAT GCTGGCTTG	1140

	CCCTGGGAGG ATGGGGCTCC TGGCCGAGAA ACTGGCCGTC ATGGGAGGCA GTGGCTGTGG	1200
5	GATTATGTGG CCATCCAACC CTCTGGATCT CCCACAGGTG AATGGCACTG TGGTGCCTACT	1260
	TGGCGAGATG AGAGGCTACG CCCCCCTTCAG CCCGGACGAG AACTCCCTGG TTCTGTTGA	1320
	AGGTTGGGGC ATGCTTCGGA ACTGGGCTGG GAGCAGGATG GTCAGCTCTT TGTCCAGTGT	1380
10	CCGGAGGAGG GACTTCCAGG AGCTGCCTGC CCTTACTCAT TTCTCCCTCC CACTGACCCC	1440
	AGGGGACGAG GTGTATTCCA CCATCCGGAA GCAGGAATAC AATGGGAAGA TCCCTCGGTT	1500
	CCGCCGCATC CGGGCGAGA GTGAGCTGTA CACCACTGAT ACTGTCTATGC AGAGTGAGTC	1560
15	AGGCTCCGGC TGGGCTGAGG GTGGGCAAGG GGGTGTGAGC ACTTAAGGTG GCAGATGGGA	1620
	TCCTGATGTT TCTGGGAGGG CTCCCTGAGG GCGGCTGGGG CCATGCAGGA AAGCAGGACC	1680
	TTGGTATAGG CCTGAGAAGT TAGGGTTGGC TGGGAGCAGA GGAACAGACA AGGTATAGCA	1740
20	GTGGGATGGG CCCAGCCCTC TTCAAGGAACA CAAACAGAGG GAGCCCCAGA CCCAGTGCAG	1800
	GGTCCCCAGG AGCCAAAGTT TATCCTCTGC TGAGTTCACG TGGAGGCAGC CCCCCAACTC	1860
	CCTCCTCATC AGGGCTCTGC CAATTGAGCA GAAGTGACAT AGGGGCCCC AGGGACCTTC	1920
25	CCCCACTCCC CAGGCATGAA GTCAATTGCTC CTGGGCGAT GACATCTTG TAGGAAGAGG	1980
	GCAAAACAGG TGTGGGTGG AGGTGCAGGG TCTAGGGCCC CTCGGGGAGT TGGACCTGAT	2040
	GTTATGAGTC CTATTCCAGA TCTGATTGTC CATGGTTTG GCAGACCCGA AGGAGGGAGG	2100
30	AGAGTGTGCA GGGTTGGAAT GGTCTCCCGG GCAAGCTTC CAGCCTTACG CCCATTGCT	2160
	TCTGTGCCCT GGCAGACCCA CAGTTCATCA AAGCCACCAT CGTGCACCAA GACCAGGCTT	2220
35	ACGATGACAA GATCTACTAC TTCTTCCGAG AGGACAATCC TGACAAGAA CCTGAGGCTC	2280
	CTCTCAATGT GTCCCGTGTG GCCCAGTTGT GCAGGGTGAA CACGGGCGTG AGGGCTGCTG	2340
	GCTACGTGTC TGTGCATGAA TAGGCCTGAG TGAGGGTGAG TTCTGTGTGT CCGTGTGCT	2400
40	GTAGAAGTTG TGTGGATGTA TGAGTGGTC TGTGTCAAGG ACTGTGGGAG CAGCTGTGTG	2460
	TGCATGGAGC ATCATGTGTC TGTGTGTGGG TAAAGGTGGC TGAGCTCCTG TGCACGTATG	2520
45	ATGGCGTGTG AGCGTGTGTA TGATGGGTG TGTGTGTGTG TGTGTGTGTG TGTTTGCT	2580
	GTGTGAATGT GCTGTGCCAC GTATGTGGGT GCGTGAGTCA GTAAATGTGT GTCTGAGTCC	2640
	GTCTGCTCTG TGGGGACCTG GCACTCTCAC CTGCCCTGAC CCTGGGCACT GCTGCCCTG	2700
50	GGCTCTGGAT CAGCCAGGCC TGCTTGCAGG AGTCTCATCT GGAGACCTGC CCTGAGTCCT	2760
	GGGGCACCCCC CGGCAGGTCC TGGCCCCCTCG CAGCCTGCCT TCCTCCCTCTG GGCCAGGTG	2820

EP 0 892 047 A2

5	TTGATATTGC TGGCAGTGGT TTCCCTGGGT GTGTGGGAA GCCCGGGCAG GTGCTGAGGG GCCTCTTCTC CCCTCTACCC TTCCAGGGG ACCAGGGTGG GGAAAGTTCA CTGTCAGTCT CCAAGTGGAA CACTTTCTG AAAGCCATGC TGGTATGCAG TGATGCTGCC ACCAACAAAGA ACTTCAACAG GCTGCAAGAC GTCTTCTGC TCCCTGACCC CAGCGGCCAG TGGAGGGACA	2880 2940 3000 3060
10	CCAGGGTCTA TGGTGTTC TCCAACCCCT GGTGAGTGGC CCTTGTCTG GGGCCGGGGC TGGCATTGGT TCAGTGTCCA GTAGGGACAG GAGGCCTTGG GCCCTGCTGA GGGCTCCCT GGTGTGGCAG GAGCAGGGC TGCAGGCTCA AGAGGCTGGG CTGTTGCTGG GTGTGGGTG GGGGGACAGC CAGTGCAGATG TATGTACTGT TGTGTGAGTG AGTCTGCCT CATGGGTGTG	3120 3180 3240 3300
15	TGTGCATGCC CTATATGCAC ACTCATGACT GCACTTGTGC CTGTTGCTGC CACCACTGC TTGTGCCGAG AGTGGACACT GGGCCAGGA GGAAGCTGCT GAAGCATCTC TCGGGGAGCT GGGTGCTATT ACACCTGCTC AGGCAGTGCC TGAGCCGAT AATTACACT TCTTAATCAC TCTCATTGAT TGAACACACG GCAGGGCGAA GTGTGGGTG TGTGTGGGA GAGTTAGGGA	3360 3420 3480 3540
20	TAGAGTGGAG GAAGCCAAGA CCCTGCTCG TGGCTCTGG GTGAGTGGGT CCCCCAGGCT GGGAAGGGGT TGGGGTCTG GCCTCCTGG GCATCAGCAC CCCACAGCCT GTGCCAGGG AGGGCTAGAG AACTGCTCAG CCTATGATGG GGTCCTCCT GCCTTGGGT TGGGTAGAGC AGATGGCCTC TAGACTCAGT GATTCTGTAA CAGGATACAA GTTTGTGGTT TTAAATTGCA	3600 3660 3720 3780
25	GCACAAAGAA ATTAGGCTGA ACTCCTCTCC TTCCCTCTCT CCATCCCTCC CCATTTTCAG TGGTGGTTGG CAACTCAGTG CCAGGCACAA GGCTGGCTG GGTGAGTGGA GGTGGATGGG TGGGTTCTGG GCCCCCATT GAGCTGGTCT CCATGTCACT GCAGGAACTA CTCAGCCGTC TGTGTGTATT CCCTCGGTGA CATTGACAAG GTCTTCCGTA CCTCCTCACT CAAGGGCTAC	3840 3900 3960 4020
30	CACTCAAGCC TTCCCAACCC GCGGCCTGGC AAGGTGAGCG TGACACCAGC CGTGGCCAG GCCAGCCCT CCTTCTGCCT CACCTCCCAC CACCCCACTG ACCTGGCCCT GCTCTCCTTG CCCAGTGCCT CCCAGACCAAG CAGCCGATAC CCACAGAGAC CTTCCAGGTG GCTGACCGTC	4080 4140 4200
35	ACCCAGAGGT GCGCAGAGG GTGGAGCCA TGGGGCTCT GAAGACGCCA TTGTTCCACT CTAAATCCA CTACCAAGAAA GTGGCGTCC ACCGCATGCA AGCCAGCCAC GGGGAGACCT TTCATGTGCT TTACCTAACT ACAGGTGAGA GGCTACCCCG GGACCCCTAG TTTGCTTTGT	4260 4320 4380
40	AAAAACCGGC ATGAAAGGTG TAAGGAATAA TGTAGTTAAC ATCTGGTTGG ATCTTTACAT	4440

EP 0 892 047 A2

	GTGGAAGGAA TAATTGAGTG ACTGGAGTTG TCAGGGGTTA ATGTGTGTGG GTGTGGAAGA	4500
5	GCCAGGCAGG GAGAGCTTCC TGGAGGGAGGT AGGGGCAAGA GGGAAAGGGG GATGGGAGAA	4550
	AAGCAAGCAC TGGGATTGAG AGCGGAAAT CTGGAGAGTC TGAGCAAAGC CAGGTGCACC	4620
	TTTGGTCCAG ATGCTGACT CAGGGAAAGA GATGGTAGGA AGAGACGTGG CAAATGAGGA	4680
10	GGAGGGCCT GAACCACAGG GATACTGCC TCTGCCAGGC AGAACATGGG AGTCAGGCC	4740
	TGGCCTGTC TTTGGGATTG TGCAGGTGAG AAGAAACATT TGAGGAGTTG ATGGGCACA	4800
	AATTAGGTAT GGGGAAGGAG TTCCAGGGGG CAGAACCTT GCCATCTCAC AGAGGACAGG	4860
15	GGCAGCTTCT CTTCTTCCCT GGAGTAGGCC CTGCTGGGG AAGCTGGTG GAATGCCGTG	4920
	GGAGATGCTC CTGCTTCTG GAAAGCCACA GGACACGGAG GAGCCAGTCC TGAGTTGGT	4980
	TTGTCGCAGC TTCCATGCC AGCTGCCCTC CTTGAGACTG GAAAGGGCCT CTAGCACCCC	5040
20	TGGGGCATT CAATTCAAGGC CCAGGCAGCCC AACCTCAGTT GTTCACATTG CCCATGTGAT	5100
	CTCCTGTTGC TGCTTCACCT TGGACTGTC TCGGCTTGG TGACCTTGTA GGAAACTGGA	5160
	ACCCCAGCAC CATTGTTGG CTCCCTGGAAG CCTTGGGAG AGGAATTTC CACAGGGCAG	5220
25	GGCCTGGTC CTGATTCCCT GCCTCTTAC TCCCTATTCA TCCCGGCTAC ACCCTTGGC	5280
	CCCCATCCTT GCTTGGCTCC AGTACTGGCT GGCACAGCTG TTGTGGTCAT CCAGGGATGG	5340
	CAGGGCACTG GGGAACAGAA GAGAGAGGTC ACACAGTGC GAACTGGGAG CAGGAGCTAG	5400
30	GACAAGGAAG GCTGGACTTG GGCCATGGAT TCCCTTCCTG CAGACTTGGG AAGTGAGCAC	5460
	ACTTGAGTGA TTAGAGAAGG TGTCTCGTT CTAAGGGCAG TGGAGGAGGC ACCATTTGG	5520
35	AGCCTGCATC ATTCTGATTT GGGCTAGATT GAAAAATAGA GCTTCTAAAG TCCCTGCAG	5580
	AGAATGGAG GCTCTCACAA CTGGGAGAAG TATTGGCTCT TTTCTGAGA ATTTGCCAA	5640
40	GGGTATGCTG TTACTGGGC TGTTTGGAA GGAGTATAGG GCATTATGTC TGTGAAGGCA	5700
	GTGGCTGGGG TGGGGCCTTA TCAGGCCAA GGACCATCTG GCCACATCTC AGAGTCACAA	5760
	GATGAGGATC ACGGATGTG AGAGGAAACA TCCTAGGCAG GCAATCATCT GACTGCTTT	5820
45	TTGGGGCAGG TGATGCCCTG GGAAATTGGG AGGGAGGGAG AGAGGGAGGT AGGCTATTCT	5880
	AGAAAATGGG AGACCAAGGTG AGGTAGGATT GGGAGGACCA GGGGTCAGGG TCCCCATTGG	5940
	TCCCTAAATTG AGAACGGAGA GAGCATTGGT CTAGGAGGCA GGCAGCTCGG TTATAAGACC	6000
50	TTGGGAACTC TTGATTAGA ATCCAAGATC CTTTTAGAT CTAGGATTAA ATAAAATTAA	6060
	GATATCCCCCT AAGATCAAAT GCAACGTGGA GTCCTGAATT GGATCCTAGA ACAGAAGAAG	6120

	GACATTTGTG GAAAAACTAG TGAAATCAA ATAAGTCTG TAGTTTGTT AATAGTAATG	6180
5	CACCAATGTC AGTTGCCTAG TTGTGACAAA TATACCGTGG TTATGTAAGA TGGTAACATT	6240
	AGGGGAACT GGAGAAGGGT AGATTGGAGC TCTCTGACT ATCTTGCAA CTTTCTGGG	6300
	AATCTAAAAT TACTCCAAA TAAAAAAAATGTTAAAGTAAATATA TTCCCTAAGA	6360
10	GTCCAGGAGG CAGGGAGTT GTAGAACGAG CTGAGTGGTT GGGTCTGAC AGATTTGGTT	6420
	CCAACTCGGT CTCTGCTGCT CACCAGCTGT GTGACCTTGA GCAAGTGGCT TAGCCTTCT	6480
	GAGCCTGATT TCCTTATCTG TGGAGTGGGG AAGATGACAG CCACCTCGCA GGGCTGTGGA	6540
15	GGGTTAAACG AGGTGATGCA TGGACACGAG CCGCAGTGAC CTTGCTGGTG TGGGCTCCT	6600
	GCTTCTGTTT TTCCCGTGCA GCCTTGGAA TGTTGGAGGC CGTATCCAGG GACCCCTGGG	6660
	CCTCCTGGGA TGGCCTCTCT GGATCAGCCT TGGAAAGGTT CAGGCTGCCCT TTAGGCTCCC	6720
20	ACATTCTTC CCAGTCACGC TCTCCTCGCC CTGCCCCACAC CAGTCCCTGIG ACCCTTGCC	6780
	GAGTTGTGAC TTCCCACCCC TCCCCGGCCT AGAGGAAAGC TGCCTGGCCC CTCAGTGGGA	6840
25	CTCCCGCCCA CTGACCCCTCT GTCCACCCATA CACAGACAGG GGCACATATCC ACAAGGTGGT	6900
	GGAACCGGGG GAGCAGGAGC ACAGCTCGC CTTCAACATC ATGGAGATCC AGCCCTTCCG	6960
	CCGGCGGGCT GCCATCCAGA CCATGTCGCT GGATGTCGAG CGGGTGAGCC TTCCCCCACT	7020
30	GCGTCCCAGTGGCTATGAG TGACTGCAGC TGAGGACAGG GCTCCTTGC ATGTGATTTG	7080
	TGTGTTCTTT TAAGAGCTTC TAGGCCTTAG GGCCTGGACA TTTAGGACTG AGTGTGGGGT	7140
	GGGGCCCGGG CCTGACCCAA TCCTGCTGTC CTTCCAGAGG AAGCTGTATG TGAGCTCCCA	7200
35	GTGGGAGGTG AGCCAGGTGC CCCTGGACCT GTGTGAGGTC TATGGGGGGG GCTGCCACGG	7260
	TTGCCTCATG TCCCGAGACC CCTACTCGG CTGGGACCAAG GCCCGCTGCA TCTCCATCTA	7320
40	CAGCTCCGAA CGGTACGTTG GCCGGATCC CTCCGTCCCT GGGACAAGGT GGGCATGGGA	7380
	CAGGGGGAGG TGTGTGCGGG CTGGAAGAGG TGGCGGTACT GGGCCTTCT TGTGGGACCT	7440
	CCTCTCTACT GGAACACTGCAC TAGGGGTAAG GATATGAGGG TCAGGTCTGC AGCCTTGTAT	7500
45	CTGCTGATCC TCTTTCGTC TTCCCACTCC AGGTCACTGTC TGCAATCCAT TAATCCAGCC	7560
	GAGCCACACA AGGAGTGTCC CAACCCAAA CCAGGTACCT GATCTGGCCC TGCTGGCGGC	7620
50	TGTGGCCCAA TGAGTGGGGT ACTGCCCTGC CCTGATTGTC CTGGTCTGAG GGAAACATGG	7680
	CCTTGTCCCTG TGGGCCCCAG GTACATGGGG CAGGATACAG TCCTGCAGAG GGAGCCCTCT	7740

EP 0 892 047 A2

	TGGTGGGATG AGCGAGACGG GAGAAAAAAG GAGGACGCTG AGGGCTGGGT TCCCCACGTT	7800
	CATTCAGAAG CCTTGTCTG GGATCCCAGT CGGTGGGAG GACACATCCT CCCCTGGAG	7860
5	CTCTTGTCC CTCCCTACGG CTGCTTCCCC ACTGCCTCCC CAGACAAGGC CCCACTGCAG	7920
	AAGGTTTCCC TGGCCCCAAA CTCTCGCTAC TACCTGAGCT GCCCCATGGA ATCCCGCCAC	7980
10	GCCACCTACT CATGGGCCA CAAGGAGAAC GTGGAGCAGA GCTGCGAACCG TGGTCACCAG	8040
	AGCCCCAACT GCATCCTGTT CATCGAGAAC CTCACGGCGC AGCAGTACGG CCACTACTTC	8100
	TGCGAGGCC AGGAGGGCTC CTACTTCCGC GAGGCTCAGC ACTGGCAGCT GCTGCCCGAG	8160
15	GACGGCATCA TGGCCGAGCA CCTGCTGGGT CATGCTGTG CCCTGGCCGC CTCCCTCTGG	8220
	CTGGGGGTGC TGCCCCACACT CACTCTTGGC TTGCTGGTCC ACTAGGGCT CCCGAGGGCTG	8280
	GGCATGCCTC AGGCTTCTGC AGCCCAGGGC ACTAGAACGT CTCACACTCA GAGCCGGCTG	8340
20	GCCCCGGAGC TCCTTGCTG CCACTTCTTC CAGGGGACAG AATAACCCAG TGGAGGATGC	8400
	CAGGCCTGGA GACGTCCAGC CGCAGGGCGC TGCTGGGCC CAGGTGGCGC ACGGATGGTG	8460
	AGGGGCTGAG AATGAGGGCA CCGACTGTGA AGCTGGGCA TCGATGACCC AAGACTTTAT	8520
25	CTTCTGGAAA ATATTTTCA GACTCCTCAA ACTTGACTAA ATGCAGCGAT GCTCCCAGCC	8580
	CAAGAGCCA TGGGTGGGG AGTGGGTTTG GATAGGAGAG CTGGGACTCC ATCTCGACCC	8640
	TGGGGCTGAG GCCTGAGTCC TTCTGGACTC TTGGTACCCA CATTGCTCC TTCCCTCTCC	8700
30	TCTCTCATGG CTGGGTGGCT GGTGTTCTG AAGACCCAGG GCTACCCCT GTCCAGCCCT	8760
	GTCTCTTGCA GCTCCCTCTC TGGTCCTGGG TCCACAGGA CAGCCGCCTT GCATGTTTAT	8820
35	TGAAGGATGT TTGCTTCCG GACGGAAGGA CGGAAAAAAGC TCTGAAAAAA AAAAAAAA	8880
	AAAAAAA	8888

40 (2) ANGABEN ZU SEQ ID NO: 42:

(i) SEQUENZKENNZEICHEN:
 (A) LÄNGE: 6622 Basenpaare
 (B) ART: Nucleotid
 (C) STRANGFORM: Einzelstrang
 (D) TOPOLOGIE: linear

45 (ii) ART DES MOLEKÜLS: Genom-DNA

50 (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..6622

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 42:

5	GATATCATGG AGATAATTAA AATGATAACC ATCTCGAAA TAAATAAGTA TTTTACTGTT	60
	TTCGTAACAG TTTTGTATAA AAAAAACCTA TAAATATGAA ATTCTTAGTC AACGTTGCC	120
	TTGTTTTAT GGTGTTACAT ATTTCTTACA TCTATGCGGA TCGATGGGGA TCCGCCCAGG	180
10	GCCACCTAAG GAGCGGACCC CGCATCTTCG CCGTCTGGAA AGGCCATGTA GGGCAGGACC	240
	GGGTGGACTT TGGCCAGACT GAGCCGCACA CGGTGCTTTT CCACGAGCCA GGCAGCTCCT	300
	CTGTGTGGGT GGGAGGACGT GGCAAGGTCT ACCTCTTGA CTTCCCCGAG GGCAAGAACG	360
15	CATCTGTGCG CACGGTGAAT ATCGGCTCCA CAAAGGGTC CTGTCTGGAT AAGCAGGACT	420
	GCGAGAACTA CATCACTCTC CTGGAGAGGC GGAGTGACGG GCTGCTGGCC TGTGGCACCA	480
20	ACGCCCGGCA CCCCAGCTGC TGGAACCTGG TGAATGGC AC TGTGGTGCCTA CTTGGCGAGA	540
	TGAGAGGCTA TGCCCCCTTC AGCCCGGACG AGAACTCCCT GGTCTGTTT GAAGGGGACG	600
	AGGTGTATTC CACCATCCGG AAGCAGGAAT ACAATGGAA GATCCCTCGG TTCCGCCGCA	660
25	TCCGGGGCGA GAGTGAGCTG TACACCAGTG ATACTGTAT GCAGAACCCA CAGTTCATCA	720
	AAGCCACCAT CGTGCACCAA GACCAGGCTT ACGATGACAA GATCTACTAC TTCTCCGAG	780
	AGGACAATCC TGACAAGAAT CCTGAGGCTC CTCTCAATGT GTCCCGTGTG GCCCAGTTGT	840
30	GCAGGGGGGA CCAGGGTGGG GAAAGTTCAC TGTCAGTCTC CAAGTGGAAC ACTTTCTGA	900
	AAGCCATGCT GGTATGCAGT GATGCTGCCA CCAACAAGAA CTTCAACAGG CTGCAAGACG	960
35	TCTTCCCTGCT CCCTGACCCC AGCGGCCAGT GGAGGGACAC CAGGGTCTAT GGTGTTTCT	1020
	CCAACCCCTG GAACTACTCA GCCGTCTGTG TGTATTCCCT CGGTGACATT GACAAGGTCT	1080
	TCCGTACCTC CTCACTCAAG GGCTACCACT CAAGCCTTCC CAACCCGCGG CCTGGCAAGT	1140
40	GCCTCCCAGA CCAGCAGCCG ATACCCACAG AGACCTTCCA GGTGGCTGAC CGTCACCCAG	1200
	AGGTGGCGCA GAGGGTGGAG CCCATGGGGC CTCTGAAGAC GCCATTGTTT CACTCTAAAT	1260
	ACCACTACCA GAAAGTGGCC GTTCACCGCA TGCAAGCCAG CCACGGGGAG ACCTTTCATG	1320
45	TGCTTTACCT AACTACAGAC AGGGGCACCA TCCACAAGGT GGTGGAACCG GGGGAGCAGG	1380
	AGCACAGCTT CGCCTTCAAC ATCATGGAGA TCCAGCCCTT CGGCCGCGCG GCTGCCATCC	1440
50	AGACCATGTC GCTGGATGCT GAGCGGAGGA AGCTGTATGT GAGCTCCCAG TGGGAGGTGA	1500
	GCCAGGTGCC CCTGGACCTG TGTGAGGTCT ATGGGGGGGG CTGCCACGGT TGCCTCATGT	1560

5	CCCGAGACCC CTACTGCGGC TGGGACCAGG GCCGCTGCAT CTCCATCTAC AGCTCCGAAC	1620
	GGTCAGTGCT GCAATCCATT AATCCAGCCG AGCCACACAA GGAGTGTCCC AACCCCAAAC	1680
10	CAGACAAGGC CCCACTGCAG AAGGTTTCCC TGGCCCCAAA CTCTCGCTAC TACCTGAGCT	1740
	GCCCCATGGA ATCCCGCAC GCCACCTACT CATGGCGCCA CAAGGAGAAC GTGGAGCAGA	1800
15	GCTGCGAACCC TGGTCACCAAG AGCCCCAACT GCATCCTGTT CATCGAGAAC CTCACGGCGC	1860
	AGCAGTACGG CCACTACTTC TGCGAGGCC AGGAGGGCTC CTACTTCCGC GAGGCTCAGC	1920
	ACTGGCAGCT GCTGCCCCAG GACGGCATCA TGCGAGGAGCA CCTGCTGGGT CATGCCTGTG	1980
20	CCCTGGCTGC CTGAATTGCA AGCTTGGAGT CGACTCTGCT GAAGAGGAGG AAATTCTCCT	2040
	TGAAGTTTCC CTGGTGTCA AAGTAAAGGA GTTTGCCACCA GACGCACCTC TGTTCACTGG	2100
	TCCGGCGTAT TAAAACACGA TACATTGTTA TTAGTACATT TATTAAGCGC TAGATTCTGT	2160
	GCGTTGTTGA TTTACAGACA ATTGTTGTAC GTATTTAAAT AATTCAATTAA ATTATATAATC	2220
	TTTAGGGTGG TATGTTAGAG CGAAAATCAA ATGATTTCA GCGTCTTTAT ATCTGAATT	2280
25	AAATATTAAA TCCTCAATAG ATTGTAAAA TAGGTTGCA TTAGTTTCAA ACAAGGGTTG	2340
	TTTTTCCGAA CCGATGGCTG GACTATCTAA TGGATTTTCG CTCAACGCCA CAAAACCTGC	2400
	CAAATCTTGT AGCAGCAATC TAGCTTGTC GATATTGTT TGTGTTTTGT TTTGTAATAA	2460
30	AGGTTGACG TCGTTCAAAA TATTATGCGC TTTTGTTATT CTTTCATCAC TGTCGTTAGT	2520
	GTACAATTGA CTCGACGTAA ACACGTTAAA TAAAGCCTGG ACATATTAA CATGGCGT	2580
35	GTTAGCTTTA TTAGGCCGAT TATCGTCGTC GTCCCAACCC TCGTCGTTAG AAGTTGCTTC	2640
	CGAAGACGAT TTGCCCCATAG CCACACGACG CCTATTAAATT GTGTCGGCTA ACACGTCCGC	2700
	GATCAAATTGTT GTAGTTGAGC TTTTTGGAAT TATTTCGTAT TGCGGGCGTT TTTGGCGGG	2760
40	TTTCAATCTA ACTGTGCCCG ATTAAATTTC AGACAAACACG TTAGAAAGCG ATGGTGCAGG	2820
	CGGTGGTAAC ATTTCAGACG GCAAATCTAC TAATGGCGGC GGTGGTGGAG CTGATGATAA	2880
45	ATCTACCATC GGTGGAGGCG CAGGGGGGC TGGCGGGCGA GGCAGGGCG GAGGTGGTGG	2940
	CGGTGATGCA GACGGCGGTT TAGGCTAAA TTGTCCTTT CAGGCAACAC AGTCGGCACC	3000
	TCAACTATTG TACTGGTTTC GGGCGTATGG TGCACCTCTCA GTACAATCTG CTCTGATGCC	3060
50	GCATAGTTAA GCCAGCCCCG ACACCCGCCA ACACCCGCTG ACGCGCCCTG ACGGGCTTGT	3120
	CTGCTCCCGG CATCCGCTTA CAGACAAGCT GTGACCGTCT CCGGGAGCTG CATGTGTCAG	3180

	AGGTTTTCAC CGTCATCACC GAAACGCGCG AGACGAAAGG GCCTCGTGAT ACGCCTATT	3240
5	TTATAGGTTA ATGTCACTGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA	3300
	AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC	3360
	ATGAGACAAT AACCTGATA AATGCTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT	3420
10	CAACATTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCCTCC TGTTTTGCT	3480
	CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT	3540
	TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTCGCCCG CGAAGAACGT	3600
15	TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC	3660
	GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC	3720
	TCACCAAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT	3780
20	GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG	3840
	AAGGAGCTAA CGCGTTTTT GCACAACATG GGGGATCATG TAACTGCCCT TGATCGTTGG	3900
	GAACCGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA	3960
25	ATGGCAACAA CGTTGCGCAA ACTATTAACG GCGAACTAC TTACTCTAGC TTCCCGGCAA	4020
	CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGAGGAC CACTTCTGCG CTCGGCCCTT	4080
	CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGGGTATC	4140
30	ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGG	4200
	AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT	4260
35	AAGCATTGGT AACTGTCAGA CCAAGTTAC TCATATATAC TTTAGATTGA TTTAAAACCTT	4320
	CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTG ATAATCTCAT GACCAAAATC	4380
	CCTTAACGTG AGTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT	4440
40	TCTTGAGATC CTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA	4500
	CCAGCGGTGG TTTGTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GTAACTGGC	4560
	TTCAGCAGAG CGCAGATACC AAATACTGTT CTTCTAGTGT AGCCGTAGTT AGGCCACAC	4620
45	TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT	4680
	GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT	4740
50	AAGGCGCAGC GGTCGGGCTG AACGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG	4800
	ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA	4860

	GGGAGAAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG	4920
5	GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA	4980
	CTTGAGCGTC GATTTTGTG ATGCTCGTCA GGGGGCGGA GCCTATGGAA AAACGCCAGC	5040
	AACGCGCCT TTTTACGGTT CCTGGCCTT TGCTGGCCTT TTGCTCACAT GTTCTTCCCT	5100
10	GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT	5160
	CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCATC CTGCACCATC	5220
	GTCTGCTCAT CCATGACCTG ACCATGCAGA GGATGATGCT CGTGACGGTT AACGCCCTGA	5280
15	ATCAGCAACG GCTTGCCGTT CAGCAGCAGC AGACCATTAA CAATCCGCAC CTGGCGAAA	5340
	CCGACATCGC AGGCTTCTGC TTCAATCAGC GTGCCGTCGG CGGTGTGCAG TTCAACCACC	5400
20	GCACGATAGA GATTGGGAT TTCCGGCCTC CACAGTTCTG GGTTTTCGAC GTTCAGACGT	5460
	AGTGTGACGC GATCGGTATA ACCACCACGC TCATCGATAA TTTCACCGCC GAAAGGCGCG	5520
	GTGCCGCTGG CGACCTGCGT TTCAACCTGC CATAAAGAAA CTGTTACCCG TAGGTAGTCA	5580
25	CGCAACTCGC CGCACATCTG AACTTCAGCC TCCAGTACAG CGCGGCTGAA ATCATCATTA	5640
	AAGCGAGTGG CAACATGGAA ATCGCTGATT TGTGTAGTCG GTTTATGCAG CAACGAGACG	5700
	TCACGGAAA TGCCGCTCAT CGGCCACATA TCCTGATCTT CCAGATAACT GCCGTCACTC	5760
30	CAACGCAGCA CCATCACCGC GAGGCGGTTT TCTCCGGCGC GTAAAAATGC GCTCAGGTCA	5820
	AATTCAAGACG GCAAACGACT GTCCCTGGCCG TAACCGACCC AGCGCCCGTT GCACCACAGA	5880
35	TGAAACGCCG AGTTAACGCC ATCAAAAATA ATTTCGCGTCT GGCCTTCCCTG TAGCCAGCTT	5940
	TCATCAACAT TAAATGTGAG CGAGTAACAA CCCGTCGGAT TCTCCGTGGG AACAAACGGC	6000
	GGATTGACCG TAATGGGATA GGTCACGTTG GTGTAGATGG GCGCATCGTA ACCGTGCATC	6060
40	TGCCAGTTG AGGGGACGAC GACAGTATCG GCCTCAGGAA GATCGCACTC CAGCCAGCTT	6120
	TCCGGCACCG CTTCTGGTGC CGGAAACCAAG GCAAAGGCC ATTTCGCCATT CAGGCTGCC	6180
45	AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCAGAAAGGG	6240
	GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGT TTTCCCAGTC ACGACGTTGT	6300
	AAAACGACGG GATCTATCAT TTTTAGCAGT GATTCTAATT GCAGCTGCTC TTTGATACAA	6360
50	CTAATTTAC GACGACGATG CGAGCTTTA TTCAACCGAG CGTGCATGTT TGCAATCGTG	6420
	CAAGCGTTAT CAATTTTCA TTATCGTATT GTTGCACATC AACAGGCTGG ACACCACGTT	6480

GAACTCGCCG CAGTTTGCG GCAAGTTGGA CCCGCCGCGC ATCCAATGCA AACTTTCCGA 6540
CATTCTGTTG CCTACGAACG ATTGATTCTT TGTCCATTGA TCGAAGCGAG TGCCTTCGAC 6600
5 TTTTTCGTGT CCAGTGTGGC TT 6622

(2) ANGABEN ZU SEQ ID NO: 43:

10

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 31 Basenpaare
(B) ART: Nucleotid
(C) STRANGFORM: Einzelstrang
15 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

20

(ix) MERKMALE:
(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..31

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 43:

CCGGATCCGC CCAGGGCCAC CTAAGGAGCG G

31

30

(2) ANGABEN ZU SEQ ID NO: 44:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 29 Basenpaare
(B) ART: Nucleotid
35 (C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 44:

CTGAATTCAG GAGCCAGGGC ACAGGCATG

29

45

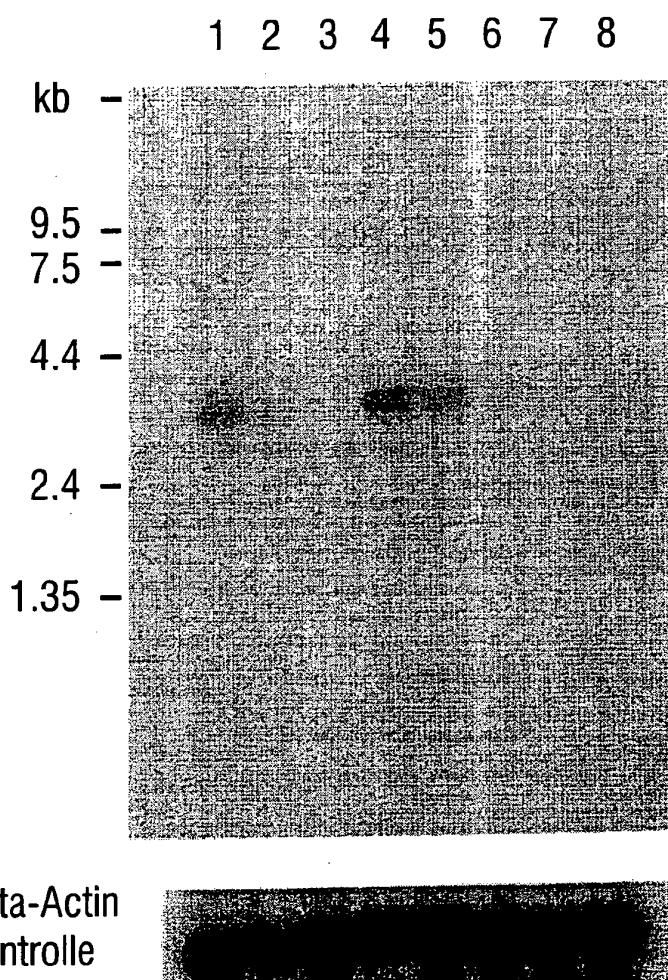
Patentansprüche

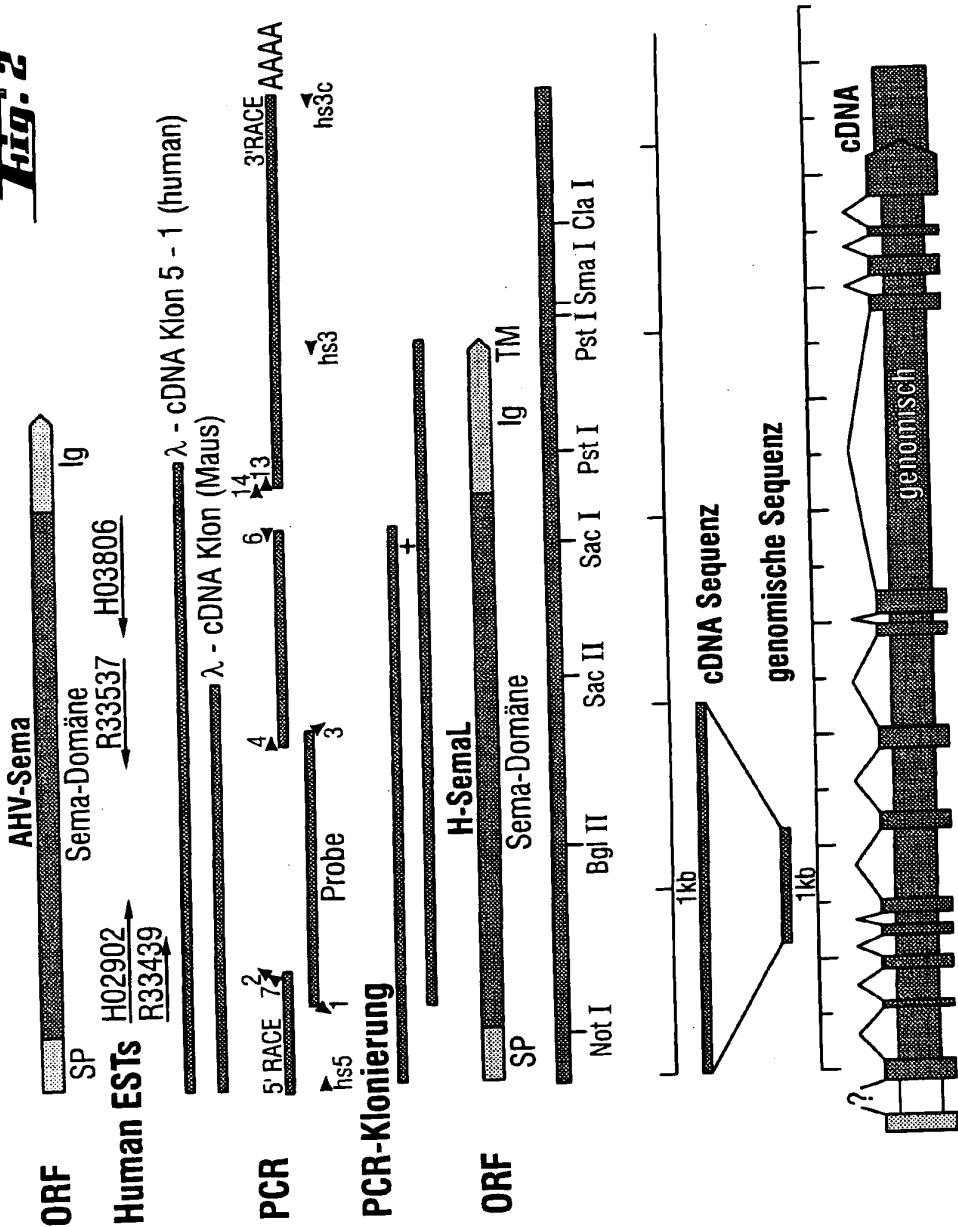
50

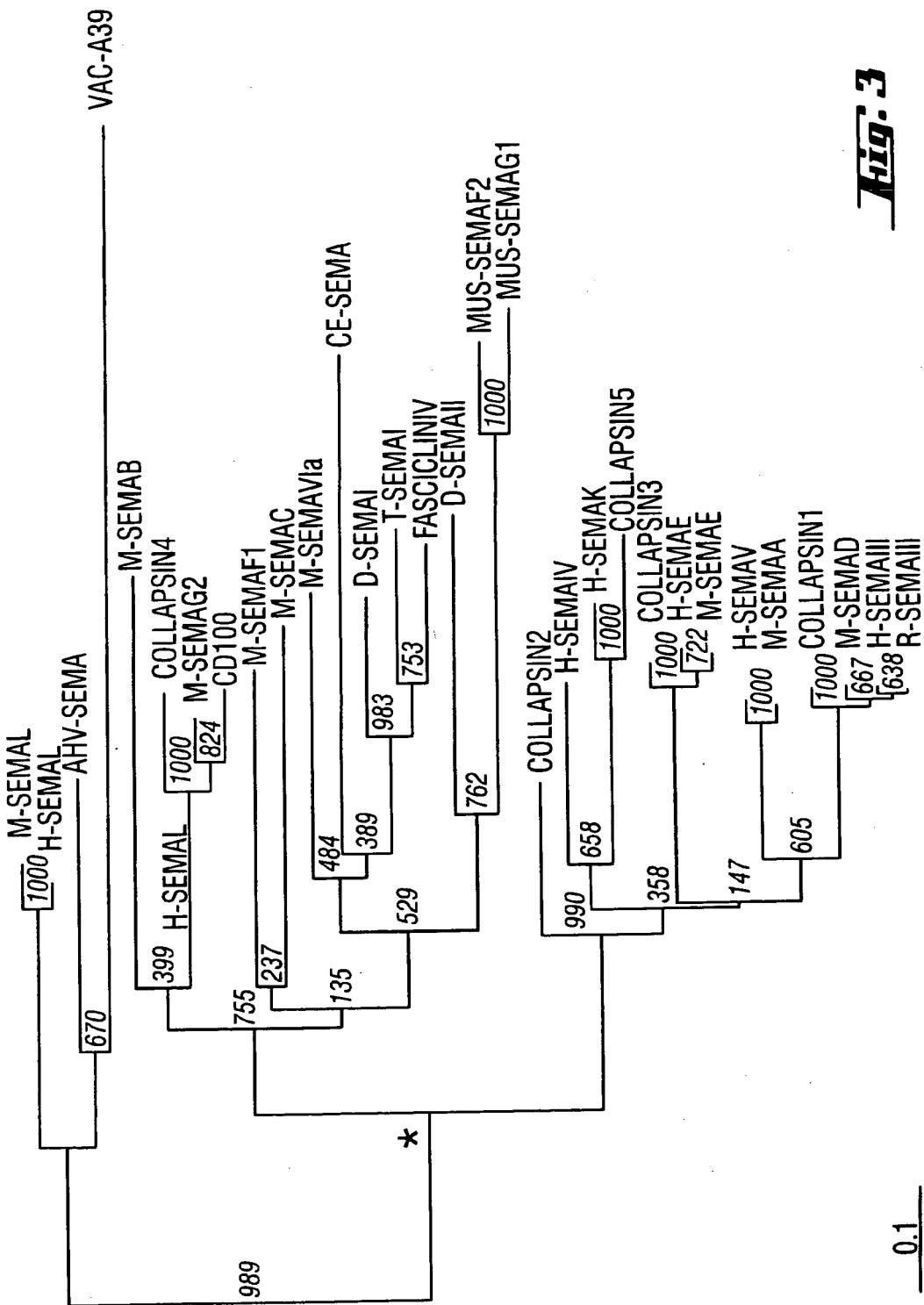
1. Semaphorin enthaltend eine charakteristische Sema-Domäne, dadurch gekennzeichnet, daß das Protein ein N-terminales Signalpeptid und im C-terminalen Bereich eine Immunglobulin-ähnliche Domäne und eine Transmembrandomäne aufweist, wobei das Semaphorin als Semaphorin vom Typ L (SemaL) bezeichnet wird und Derivate des Semaphorins vom Typ L
- 55 2. Semaphorin nach Anspruch 1, wobei das Protein (humanes Semaphorin vom Typ L (H-SemaL)) die Aminosäuresequenz SEQ ID NO. 3 hat.

3. Semaphorin nach einem oder mehreren der Ansprüche 1 und 2, wobei das Protein im Bereich der Sema-Domäne eine Aminosäureidentität von mindestens 40 % im Bezug auf die Sema-Domäne von H-SemaL aufweist.
4. Semaphorin nach einem oder mehreren der Ansprüche 1 bis 2, wobei das Protein die partielle Aminosäuresequenz SEQ ID NO. 4 enthält (murines Semaphorin (M-SemaL)).
5. Nukleinsäure enthaltend eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L gemäß einem oder mehreren der Ansprüche 1 bis 4 kodiert sowie Derivate derselben.
10. Nukleinsäure nach Anspruch 5, wobei besagte Nukleinsäure-Sequenz ein Semaphorin L Gen ist.
7. Nukleinsäure nach einem oder mehreren der Ansprüche 5 und 6, wobei besagte Nukleinsäure-Sequenz das Gen von H-SemaL enthält.
15. Nukleinsäure nach Anspruch 5, wobei besagte Nukleinsäure-Sequenz die cDNA eines Semaphorins vom Typ L enthält.
9. Nukleinsäure nach Anspruch 8, wobei die cDNA die cDNA von H-SemaL ist.
20. 10. Nukleinsäure nach Anspruch 8, wobei die cDNA die cDNA von M-SemaL ist.
11. Verfahren zur Herstellung eines Semaphorins vom Typ L gemäß einem oder mehreren der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L oder ein Derivat derselben kodiert, in einen Expressionsvektor kloniert und exprimiert wird.
25. 12. Verfahren gemäß Anspruch 11, wobei für die Expression eine eukaryotische Zelle verwendet wird.
13. Verwendung eines Semaphorins vom Typ L oder eines Derivats desselben oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder eines Derivats derselben zur Herstellung eines Arzneimittels, das zur Behandlung oder Prävention von immunologischen Erkrankungen verwendet werden kann.
30. 14. Verwendung einer Nukleinsäure-Sequenz oder eines Derivats derselben nach Anspruch 13 in der Gentherapie.
15. Verwendung eines Semaphorins vom Typ L oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert in einem Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen.
35. 16. Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen, dadurch gekennzeichnet, daß ein Semaphorin vom Typ L unter definierten Bedingungen mit einem zu untersuchenden Wirkstoff inkubiert wird, parallel ein zweiter Ansatz ohne den zu untersuchenden Wirkstoff, aber unter ansonsten gleichen Bedingungen durchgeführt wird und dann die inhibierende bzw. aktivierende Wirkung des zu untersuchenden Wirkstoffs bestimmt wird.
40. 17. Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen, dadurch gekennzeichnet, daß eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert unter definierten Bedingungen und in Gegenwart eines zu untersuchenden Wirkstoffs exprimiert wird und das Ausmaß der Expression bestimmt wird.
45. 18. Verfahren zur Herstellung einer Nukleinsäure, die für ein Semaphorin vom Typ L kodiert, wobei diese Nukleinsäure mit Hilfe der Polymerase Ketten Reaktion unter Verwendung von spezifischen Primern amplifiziert wird.
19. Semaphorin Antikörper, dadurch gekennzeichnet, daß er entweder das Epitop von H-SemaL, das den Aminosäuren 179-378 in SED ID NO: 4 entspricht, erkennt oder daß er das Epitop von H-SemaL, das den Aminosäuren 480-666 in SED ID NO: 4 entspricht, erkennt.
50. 20. Verfahren zur Herstellung eines Semaphorin Antikörpers nach Anspruch 19, wobei die Epitope als Fusionsproteine mit einem Epitop-Tag exprimiert und über dieses Epitop-Tag aufgereinigt werden und die aufgereinigten Fusionsproteine zur Immunisierung verwendet werden.

Fig. 1



Hig. 2



01

Abb. 4a (U937)

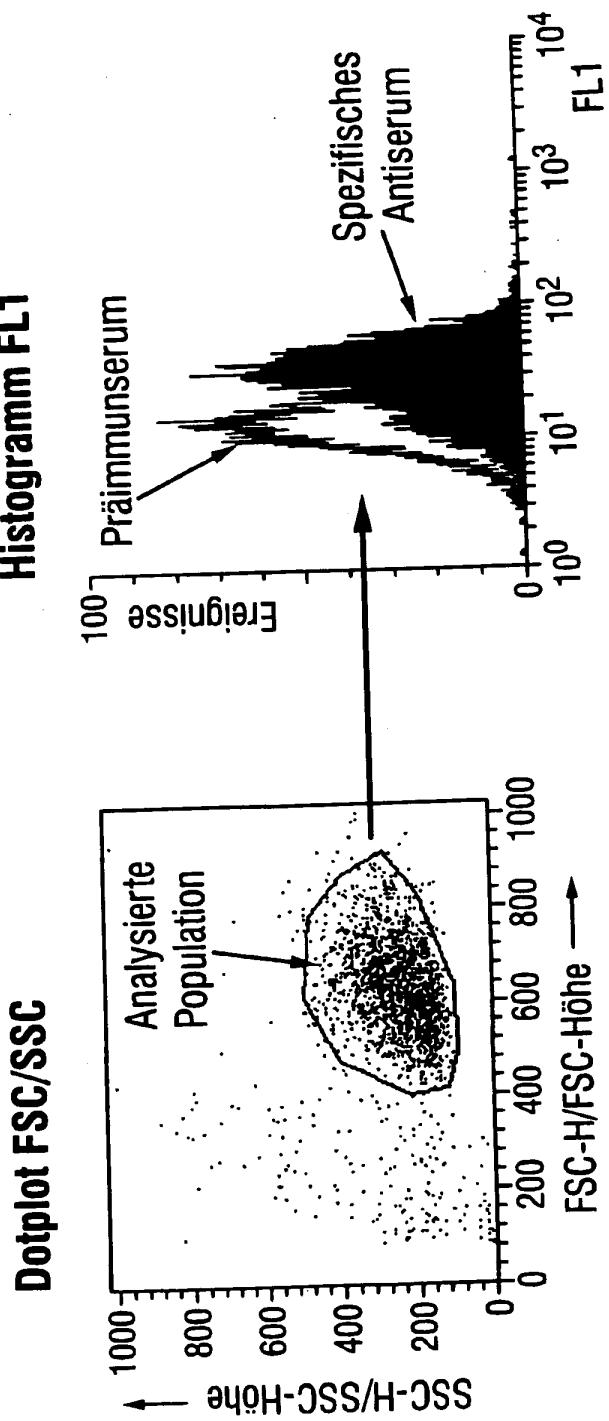


Abb. 4b (THP-1)

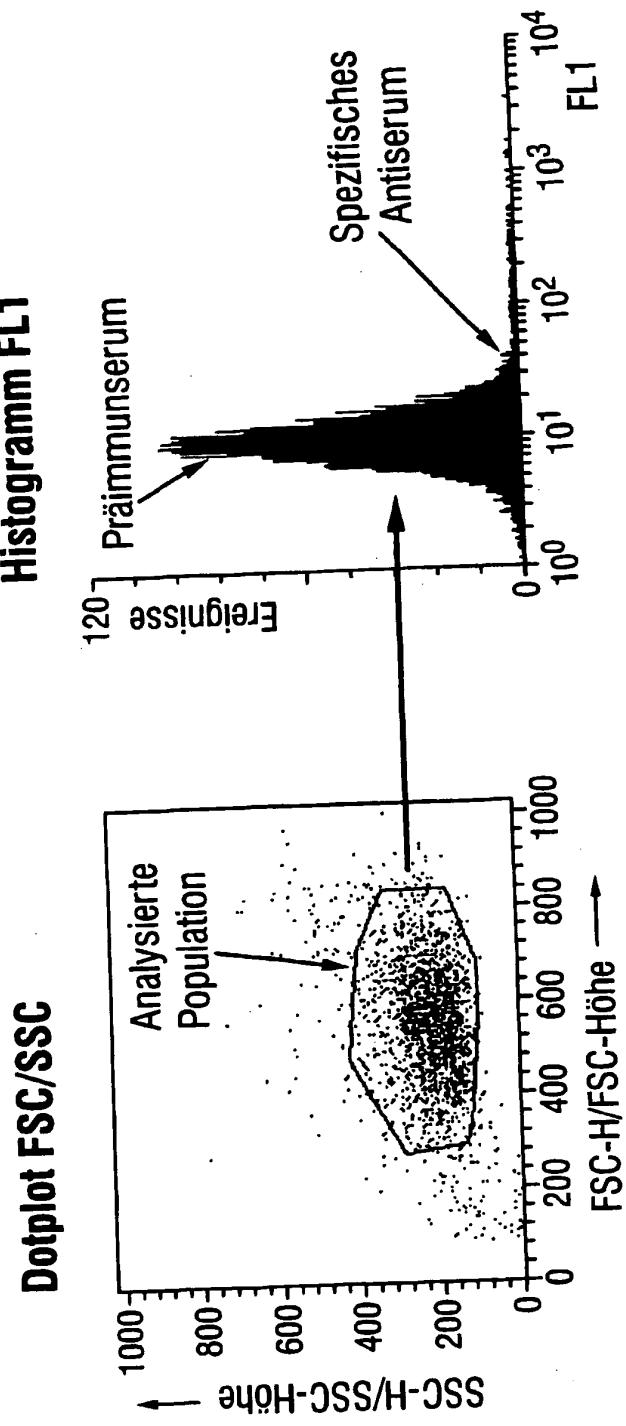


Fig. 4c (K562)

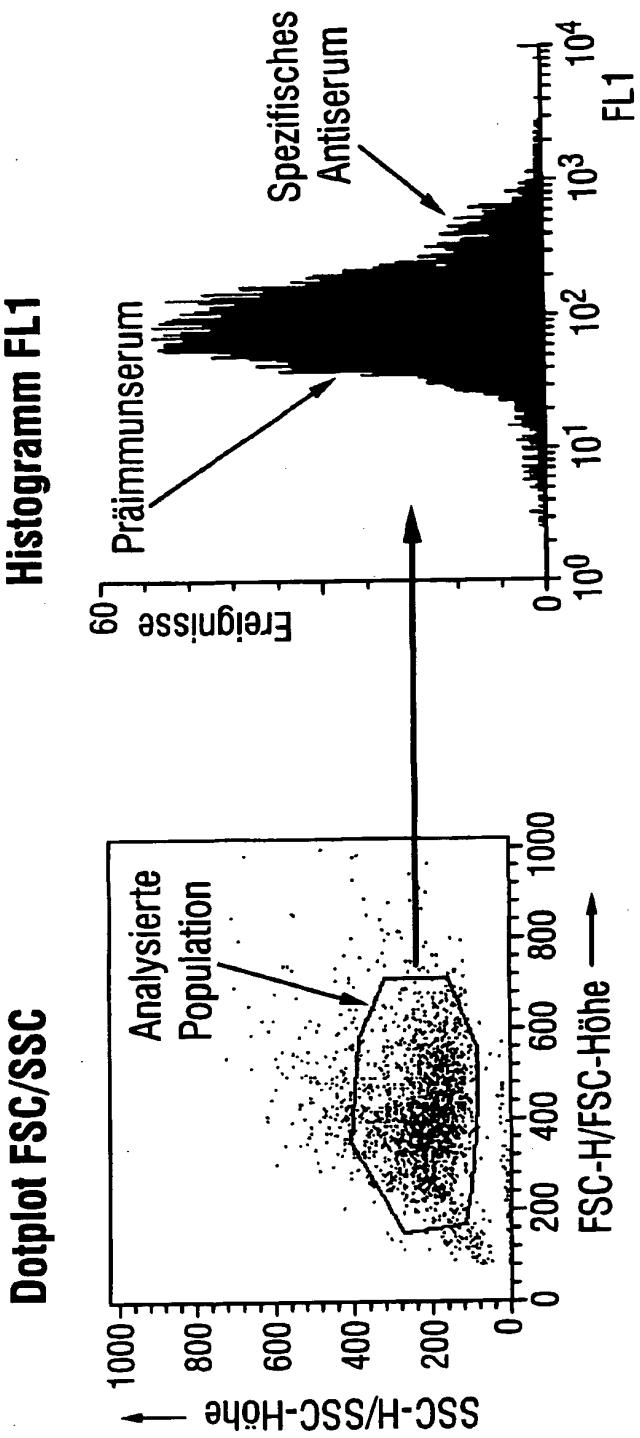


Fig. 4d (L-428)

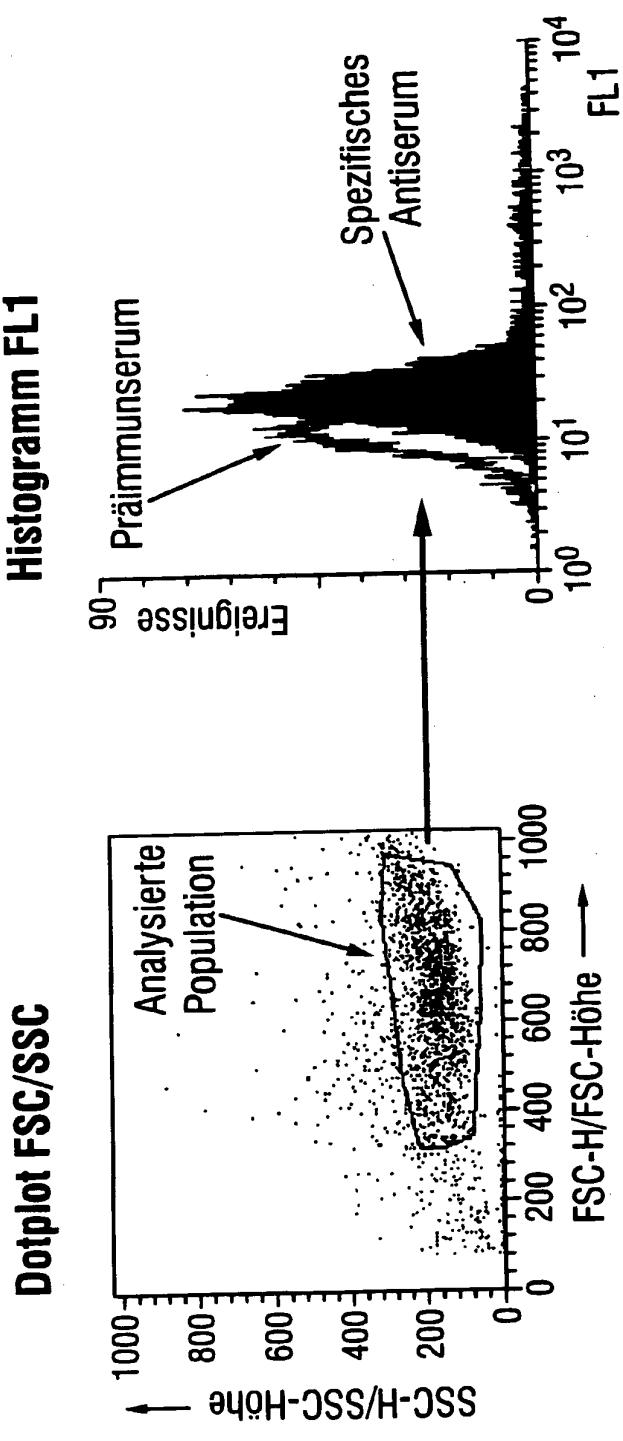


Fig. 4e (Jurkat)

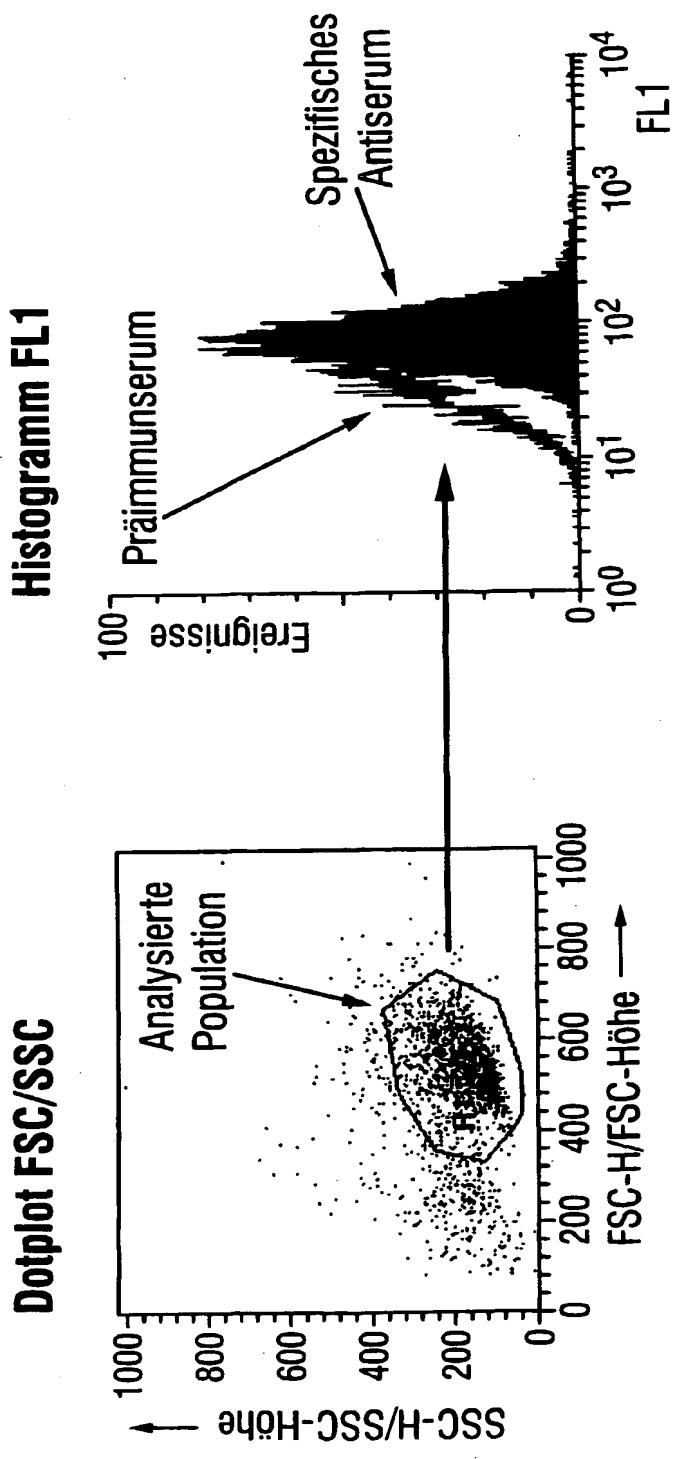


Fig. 4f (Daudi)

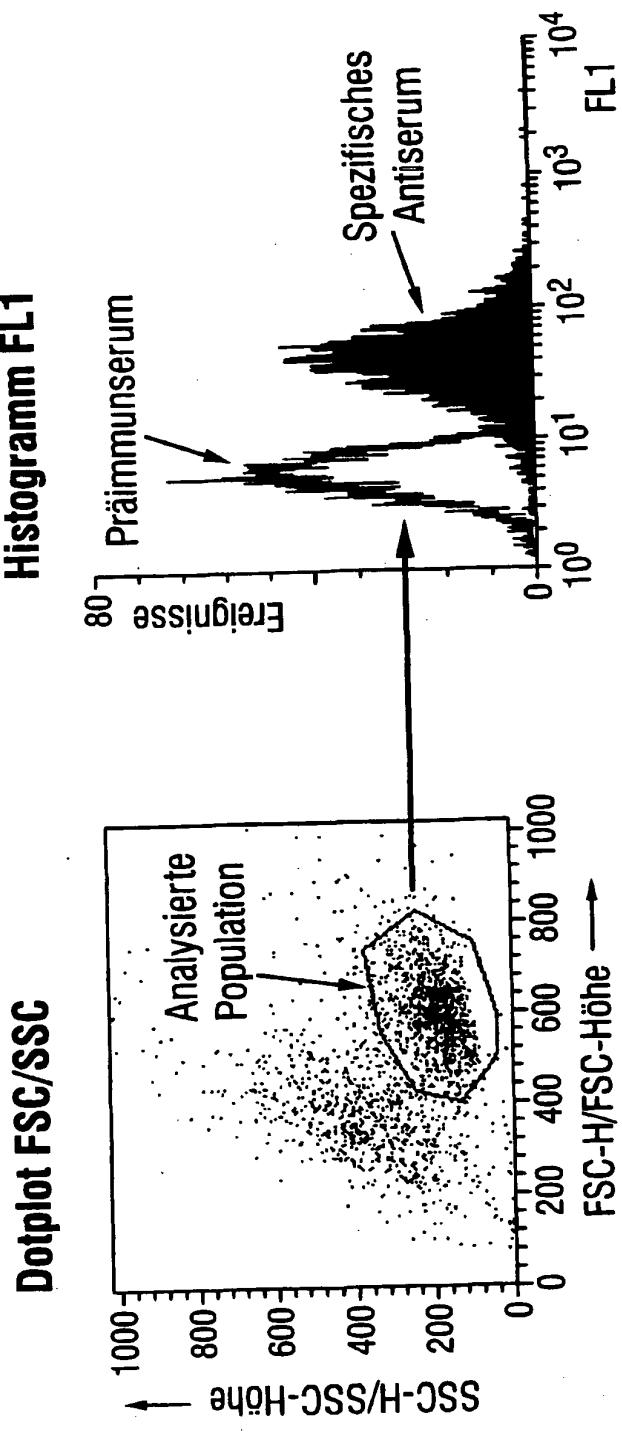


Fig. 4g (LCL EBV-Transformierte B-Zellen)

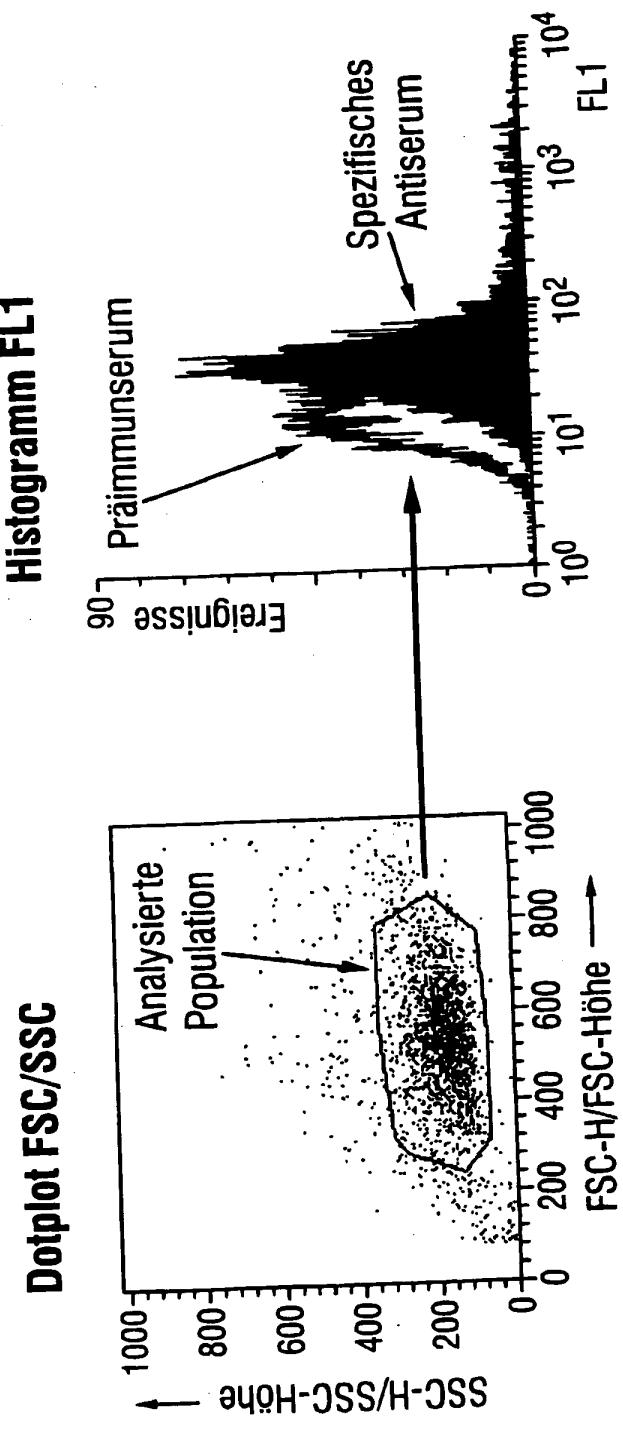


Fig. 4h (Jiyoje)

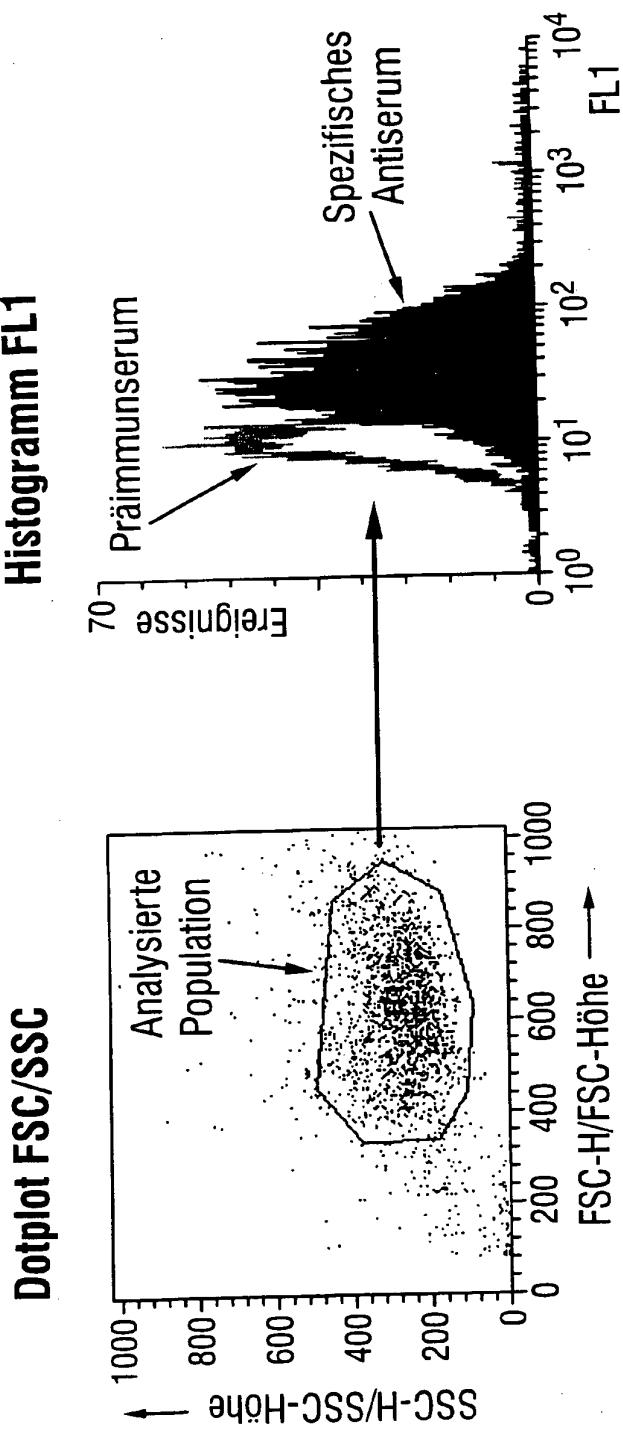


Abb. 4 i (CBL-Mix57)

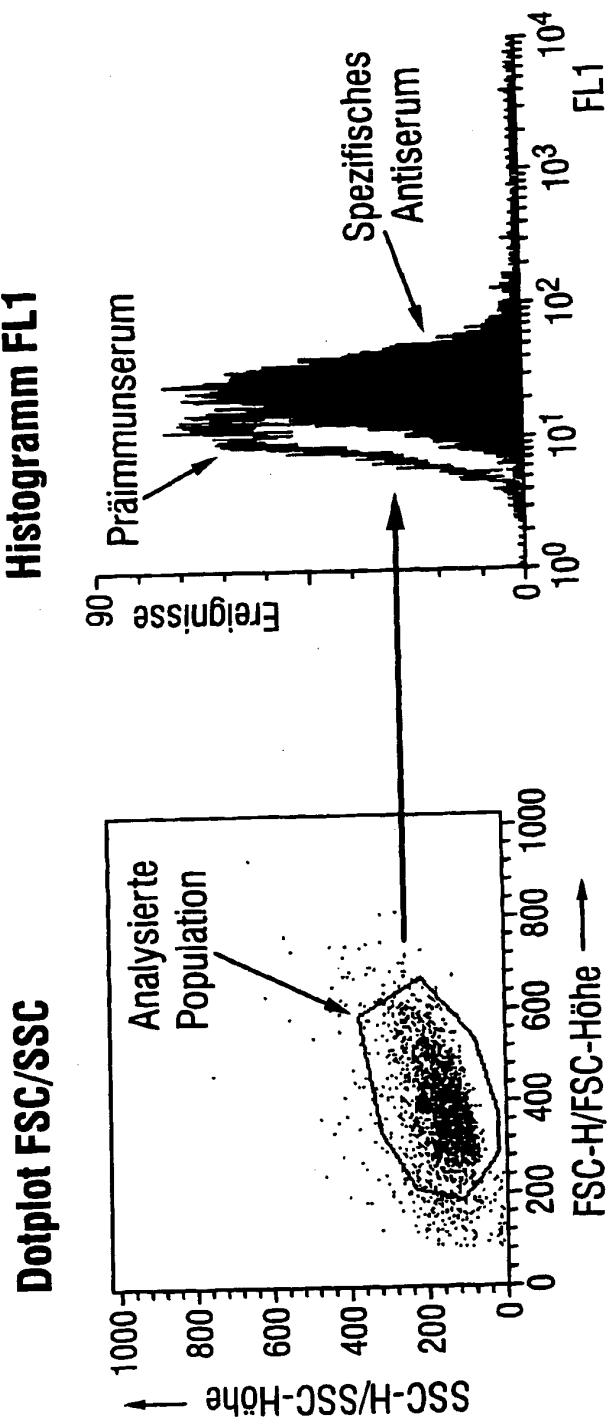


Fig. 4k (CBL-Mix59)

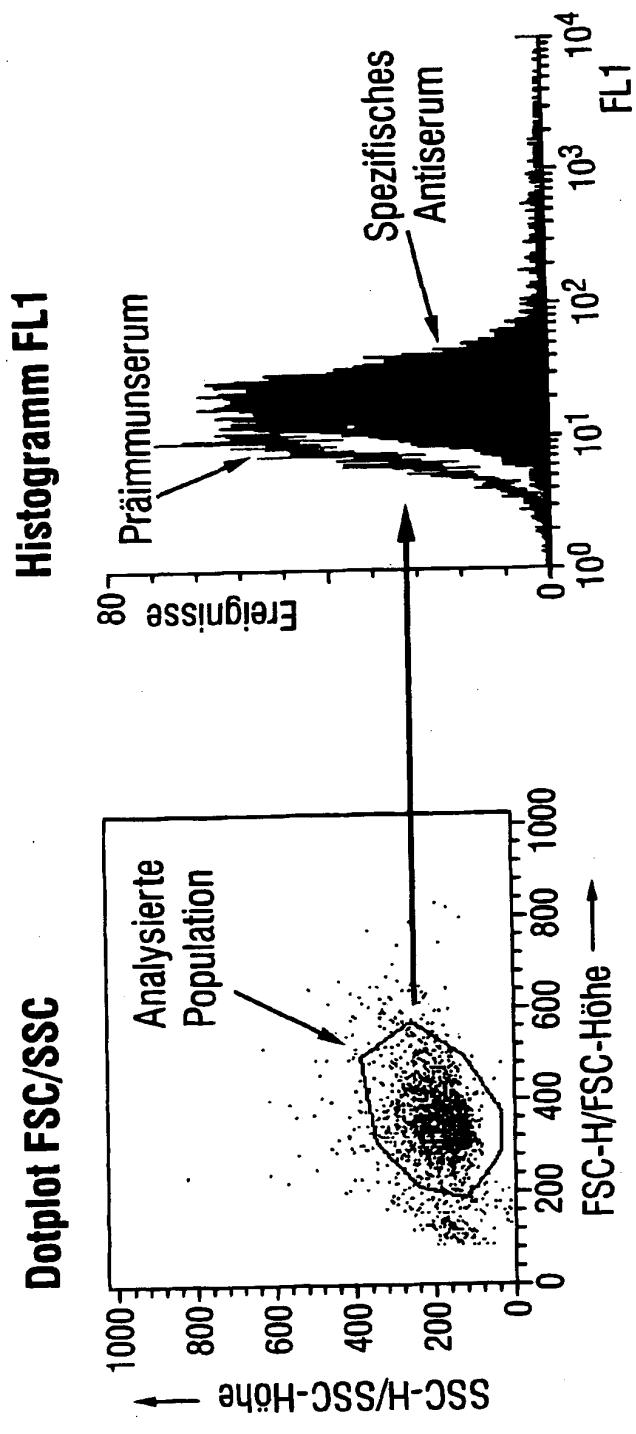
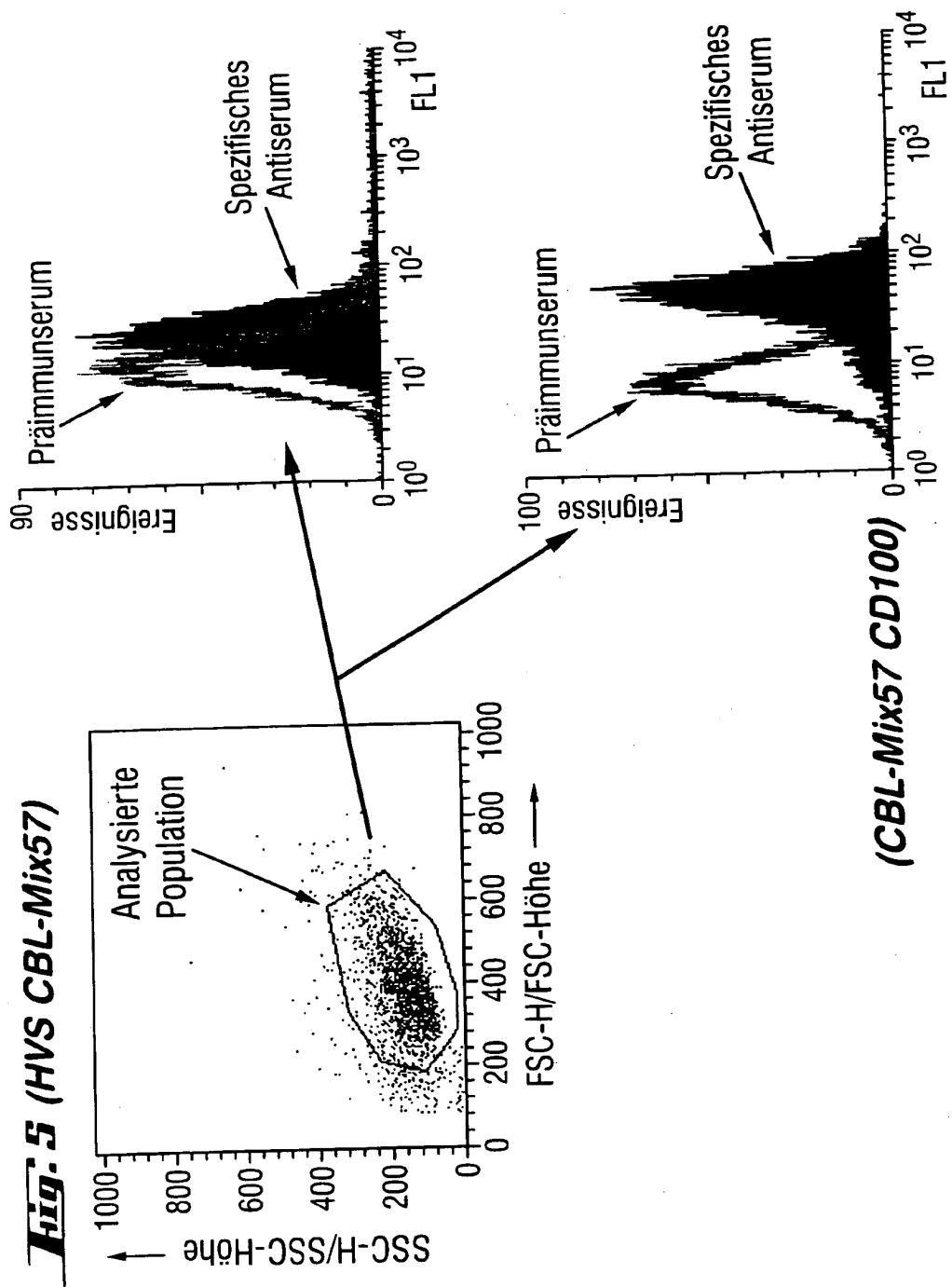


Abb. 5 (HVS CBL-Mix57)



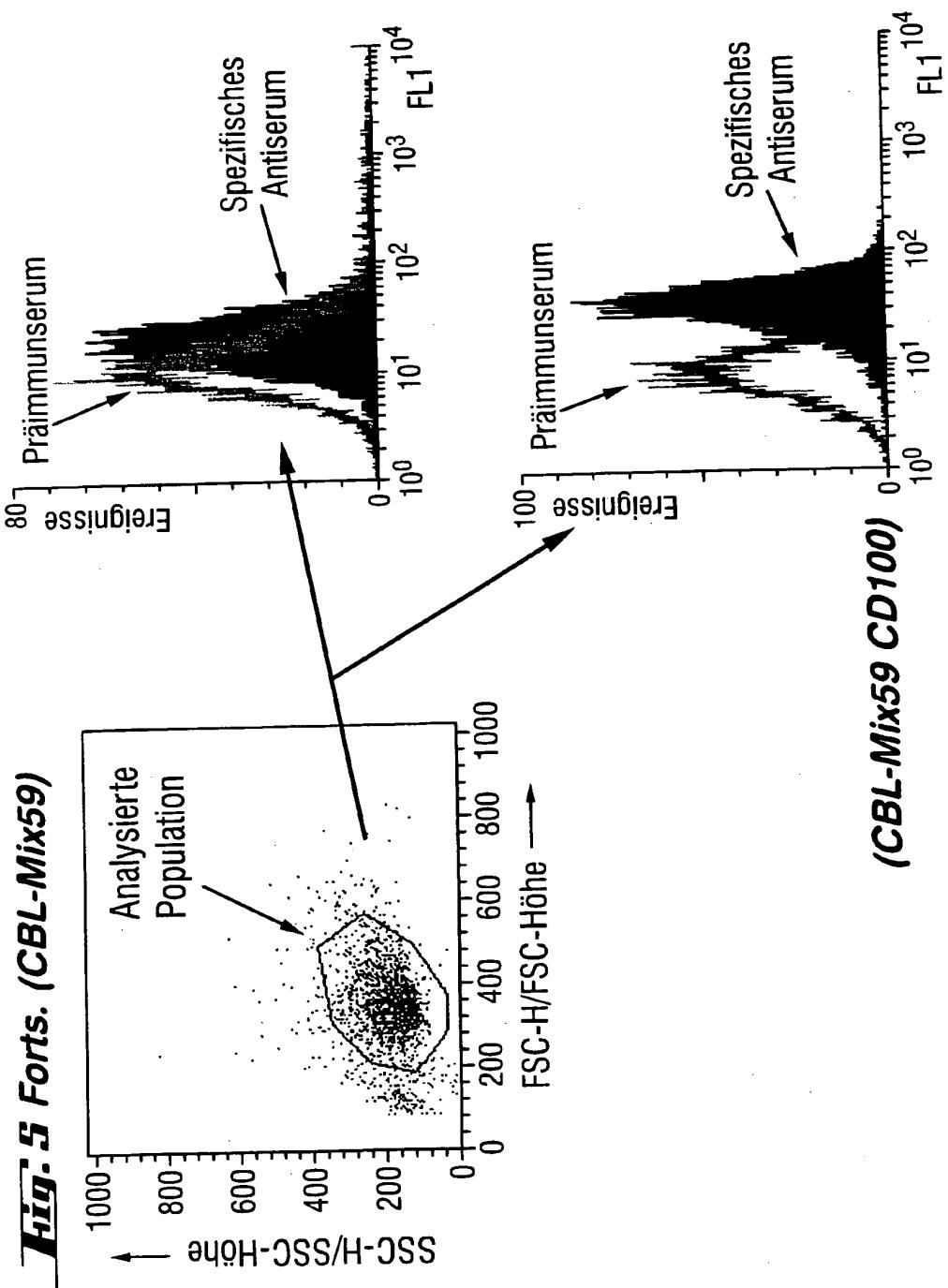


Fig. 6

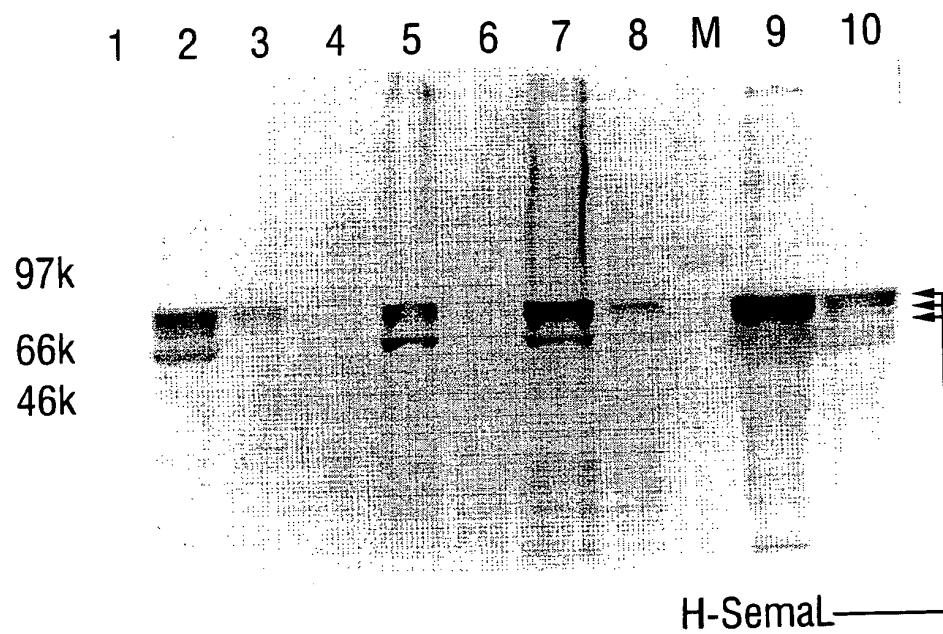
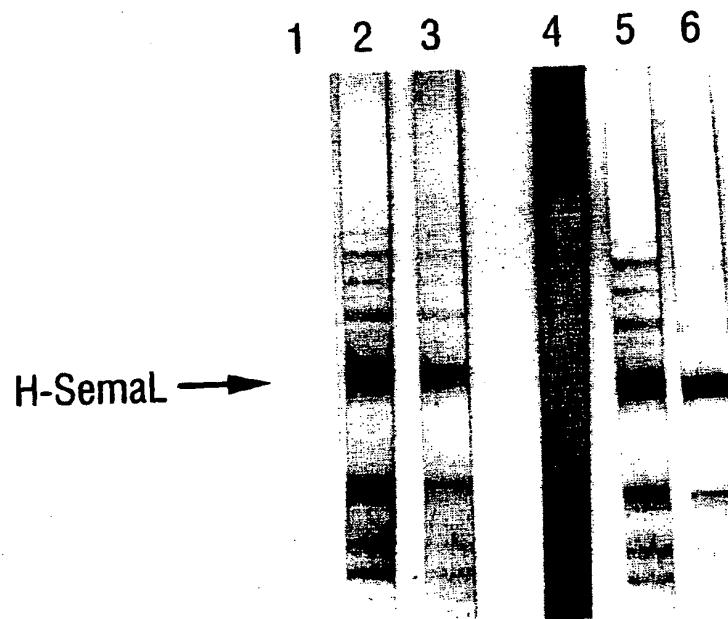


Fig. 2



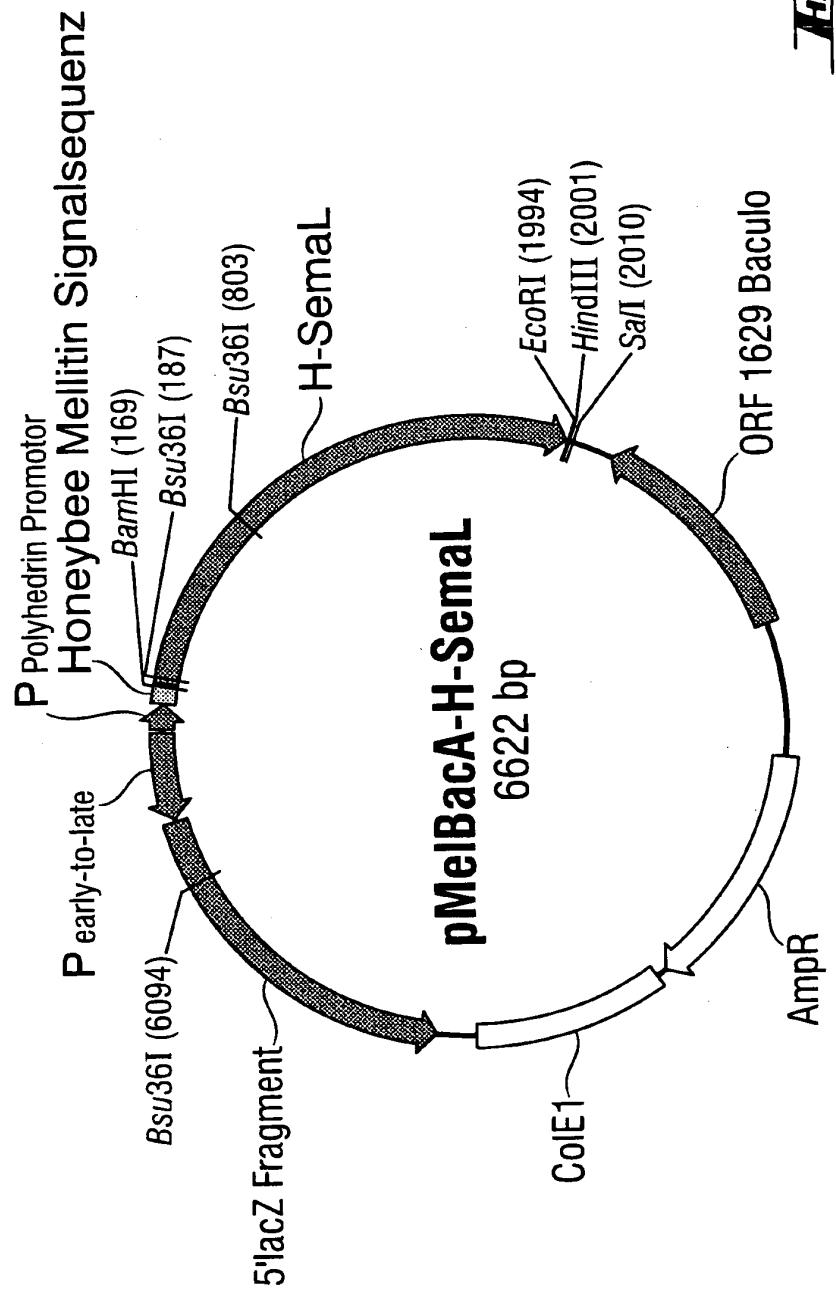


Fig. 8



(19) Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) EP 0 892 047 A3

(12) **EUROPÄISCHE PATENTANMELDUNG**

(88) Veröffentlichungstag A3:
08.03.2000 Patentblatt 2000/10

(51) Int. Cl.⁷: C12N 15/12, C07K 14/47,
A61K 48/00, G01N 33/50

(43) Veröffentlichungstag A2:
20.01.1999 Patentblatt 1999/03

(21) Anmeldenummer: 98112470.4

(22) Anmeldetag: 06.07.1998

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE

Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 09.07.1997 DE 19729211
11.02.1998 DE 19805371

(71) Anmelder:
Hoechst Marion Roussel Deutschland GmbH
65929 Frankfurt am Main (DE)

(72) Erfinder:
• Fleckenstein, Bernhard, Prof. Dr.
91369 Wiesenthau (DE)
• Ensser, Armin, Dr.
90419 Nürnberg (DE)

(54) **Humanes und murines Semaphorin L**

(57) Humanes Semaphorin L(H-SemaL) und korrespondierende Semaphorine in anderen Spezies.

Gegenstand der Erfindung sind neuen Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Verwendung derselben.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunoglobulin-ähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

EP 0 892 047 A3

Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

EINSCHLÄGIGE DOKUMENTE												
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrift Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.6)									
D, X	ENSSER A ET AL: "Alcephaline herpesvirus type 1 has a semaphorin-like gene" JOURNAL OF GENERAL VIROLOGY., Bd. 76, 1995, Seiten 1063-1067, XP002113888 SOCIETY FOR GENERAL MICROBIOLOGY, READING., GB ISSN: 0022-1317 * Abbildung 2 *	1,3,5,6, 8,11	C12N15/12 C07K14/47 A61K48/00 G01N33/50									
D, A	DATABASE EMBL 'Online! Accession Number H03806, 22. Juni 1995 (1995-06-22) HILLIER L ET AL: "yj39f01.r1 Homo sapiens cDNA clone 151129 5' similar to SP:A49069 A49069 COLLAPSIN - ;" XP002125533 * 94% Identität (420 BP) zwischen SEQ ID 1 und H03806 *	5										
A	DATABASE EMBL 'Online! accession number AA260340, 19. März 1997 (1997-03-19) MARRA M ET AL: "va91b02.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK ACCESSION NUMBER L26081. ;" XP002125534 * 99.7% Identität (347 BP) zwischen SEQ ID 1 und H03806 *	4	RECHERCHIERTE SACHGEBiete (Int.Cl.6) C12N C07K A61K G01N									
<p>Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt</p> <table border="1"> <tr> <td>Recherchenort DEN HAAG</td> <td>Abschlußdatum der Recherche 10. Januar 2000</td> <td>Prüfer Lejeune, R</td> </tr> <tr> <td colspan="3"> KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur </td> </tr> <tr> <td colspan="3"> T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument </td> </tr> </table>				Recherchenort DEN HAAG	Abschlußdatum der Recherche 10. Januar 2000	Prüfer Lejeune, R	KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur			T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument		
Recherchenort DEN HAAG	Abschlußdatum der Recherche 10. Januar 2000	Prüfer Lejeune, R										
KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur												
T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument												

Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

EINSCHLÄGIGE DOKUMENTE			
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betritt Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.6)
T	<p>YAMADA A ET AL: "Molecular cloning of a glycosylphosphatidylinositol-anchored molecule CDW108" JOURNAL OF IMMUNOLOGY., Bd. 162, 1. April 1999 (1999-04-01), Seiten 4094-4100, XP002123609 THE WILLIAMS AND WILKINS CO. BALTIMORE., US ISSN: 0022-1767 * 99.8% identität zwischen SEQ ID 1 und CDW108 (Abbildung 1) * * Seite 4094, Spalte 2, Absatz 2 *</p>	1-3,5-9, 11,12, 18-20	
T	<p>LANGE C ET AL: "New eukaryotic semaphorins with close homology to semaphorins of viruses" GENOMICS, Bd. 51, 1. August 1998 (1998-08-01), Seiten 340-350, XP002113887 SAN DIEGO., US ISSN: 0888-7543 * 99.8% identität zwischen SEQ ID 1 und H-Sema-L (Abbildung 3) * -& DATABASE EMBL 'Online! Accession Number AF030699, 8. September 1998 (1998-09-08) LANGE C ET AL: "Mus musculus semaphorin L (Semal) mRNA, partial cds." XP002125537 * M-SemaL ist identisch zu SEQ ID 1 *</p>	1-12,18 -/-	RECHERCHIERTE SACHGEBiete (Int.Cl.6)
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt			
Recherchenort	Abschlußdatum der Recherche	Prüfer	
DEN HAAG	10. Januar 2000	Lejeune, R	
KATEGORIE DER GENANNTEN DOKUMENTE		T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument	
X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichttechnische Offenbarung P : Zwischenliteratur			



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

EINSCHLÄGIGE DOKUMENTE					
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrift Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.6)		
T	XU X ET AL: "Human semaphorin K1 is glycosylphosphatidylinositol-linked and defines a new subfamily of viral-related semaphorins" JOURNAL OF BIOLOGICAL CHEMISTRY., Bd. 273, Nr. 35, 28. August 1998 (1998-08-28), Seiten 22428-22434, XP002113886 AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD., US ISSN: 0021-9258 * 99.1% Identität zwischen SEQ ID 1 und Sema K1 (Abbildung 1) *	1-3,5-9, 11,12,18			
E	WO 99 45114 A (ZYMOGENETICS INC) 10. September 1999 (1999-09-10) * 99.8% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09945114) * * Beispiel 4 *	1-3,5-9, 11,12, 18-20			
E	WO 99 38885 A (SMITHKLINE BEECHAM PLC) 5. August 1999 (1999-08-05) * 99.9% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09938885 *	1-3,5-9, 11,12,18	RECHERCHIERTE SACHGEBIETE (Int.Cl.6)		
E	WO 99 58676 A (SPRIGGS MELANIE K ; IMMUNEX CORP (US)) 18. November 1999 (1999-11-18) * 99.95% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09958676 *	1-3,5-9, 11,12,18			
-/-					
Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt					
Recherchenort	Abschlußdatum der Recherche	Prüfer			
DEN HAAG	10. Januar 2000	Lejeune, R			
KATEGORIE DER GENANNTEN DOKUMENTE					
X : von besonderer Bedeutung allein betrachtet	T : der Erläuterung zugrunde liegende Theorien oder Grundsätze				
Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie	E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist				
A : technologischer Hintergrund	D : in der Anmeldung angeführtes Dokument				
O : nichtschriftliche Offenbarung	L : aus anderen Gründen angeführtes Dokument				
P : Zwischenliteratur	& : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument				



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

EINSCHLÄGIGE DOKUMENTE												
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betrifft Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.)									
T	<p>DATABASE EMBL 'Online! Accession Number AF176670, 7. September 1999 (1999-09-07) MINE T ET AL: "Mus musculus GPI-anchored membrane protein CDW108 precursor (CDW108) mRNA, complete cds." XP002125535 * 99.8% identität zwischen SEQ ID 2 und AF176670 *</p> <p>-----</p> <p>T</p> <p>DATABASE EMBL 'Online! Accession number ab017532, 15. März 1999 (1999-03-15) TAKAHASHI H ET AL: "Mus musculus mRNA for msemKip, complete cds." XP002125536 * 99.8% identität zwischen SEQ ID 2 und AB017532 *</p> <p>-----</p>	1,3-6,8, 10-12,18										
			RECHERCHIERTE SACHGEBiete (Int.Cl.)									
<p>Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 33%;">Recherchenort</td> <td style="width: 33%;">Abschlußdatum der Recherche</td> <td style="width: 34%;">Prüfer</td> </tr> <tr> <td>DEN HAAG</td> <td>10. Januar 2000</td> <td>Lejeune, R</td> </tr> <tr> <td colspan="3"> KATEGORIE DER GENANNTEN DOKUMENTE <p>X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur</p> <p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument</p> </td> </tr> </table>				Recherchenort	Abschlußdatum der Recherche	Prüfer	DEN HAAG	10. Januar 2000	Lejeune, R	KATEGORIE DER GENANNTEN DOKUMENTE <p>X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur</p> <p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument</p>		
Recherchenort	Abschlußdatum der Recherche	Prüfer										
DEN HAAG	10. Januar 2000	Lejeune, R										
KATEGORIE DER GENANNTEN DOKUMENTE <p>X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur</p> <p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument</p>												

**ANHANG ZUM EUROPÄISCHEN RECHERCHENBERICHT
ÜBER DIE EUROPÄISCHE PATENTANMELDUNG NR.**

EP 98 11 2470

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten europäischen Recherchenbericht angeführten Patentdokumente angegeben.
Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am
Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

10-01-2000

Im Recherchenbericht angeführtes Patentdokument		Datum der Veröffentlichung		Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9945114	A	10-09-1999	AU	2982799 A	20-09-1999
WO 9938885	A	05-08-1999		KEINE	
WO 9958676	A	18-11-1999		KEINE	

EPO FORM P0461

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.